NASA/TM-2003-212172

Evaluation of a Dispatcher's Route Optimization Decision Aid to Avoid Aviation Weather Hazards

Michael C. Dorneich, Olu Olofinboba, Steve Pratt, Dannielle Osborne, and Thea Feyereisen
Honeywell, Inc., Minneapolis, Minnesota

Kara Latorella
NASA Langley Research Center, Hampton Virginia

April 2003

The NASA STI Program Office . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program Office plays a key part in helping NASA maintain this important role.

The NASA STI Program Office is operated by Langley Research Center, the lead center for NASA's scientific and technical information. The NASA STI Program Office provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program Office is also NASA's institutional mechanism for disseminating the results of its research and development activities. These results are published by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers, but having less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services that complement the STI Program Office's diverse offerings include creating custom thesauri, building customized databases, organizing and publishing research results ... even providing videos.

For more information about the NASA STI Program Office, see the following:

- Access the NASA STI Program Home Page at ***http://www.sti.nasa.gov***

- E-mail your question via the Internet to help@sti.nasa.gov

- Fax your question to the NASA STI Help Desk at (301) 621-0134

- Phone the NASA STI Help Desk at (301) 621-0390

- Write to:
  NASA STI Help Desk
  NASA Center for AeroSpace Information
  7121 Standard Drive
  Hanover, MD 21076-1320

NASA/TM-2003-212172

Evaluation of a Dispatcher's Route Optimization Decision Aid to Avoid Aviation Weather Hazards

Michael C. Dorneich, Olu Olofinboba, Steve Pratt, Dannielle Osborne, and Thea Feyereisen
Honeywell, Inc., Minneapolis, Minnesota

Kara Latorella
NASA Langley Research Center, Hampton Virginia

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23681-2199

April 2003

Available from:

NASA Center for AeroSpace Information (CASI)
7121 Standard Drive
Hanover, MD 21076-1320
(301) 621-0390

National Technical Information Service (NTIS)
5285 Port Royal Road
Springfield, VA 22161-2171
(703) 605-6000

Table of Contents

List of Abbreviations & Acronyms

ADDS	Aviation Digital Data Service
ADF	Airline Dispatchers Federation
AIRMET	AIRman's METeorological Information
ANOVA	Analysis of Variance
ATC	Air Traffic Control
AWIN	Aviation Weather Information
CCFP	Collaborative Convective Forecast Product
DBz	decibels of Z (reflectivity)
DOF	Degree of Freedom
DUATS	Direct User Access Terminal System flight planning system
ERAU	Embry-Riddle Aeronautical University
ETA	Estimated Time of Ariival
FSS	Flight Service Station
GA	General Aviation
IFR	Instrument Flight Rules
METAR	METeorological Aviation Report
MOCK	Mock Weather Information Service
NASA	National Aeronautical & Space Administration
NCAR	National Center for Atmospheric Research
PIREPs	Pilot Reports
SA	Situation Awareness
SIGMET	SIGnificant METeorological Information
TAF	Terminal Aerodrome Forecasts
TLX	Task Load Index
UTC	Universal Time
WSI	Weather Services International
3D	Three Dimensional

List of Figures

List of Tables

1. Executive Summary

This document describes the results and analysis of the formal evaluation plan for the Honeywell software tool developed under the NASA AWIN (Aviation Weather Information) "Weather Avoidance using Route Optimization as a Decision Aid" project. The software tool aims to provide airline dispatchers with a decision aid for selecting optimal routes that avoid weather and other hazards. This evaluation compares and contrasts route selection performance with the AWIN tool to that of subjects using a more traditional dispatcher environment. The evaluation assesses gains in safety, in fuel efficiency of planned routes, and in time efficiency in the pre-flight dispatch process through the use of the AWIN decision aid. In addition, we are interested in how this AWIN tool affects constructs that can be related to performance. The construct of Situation Awareness (SA), workload, trust in an information system, and operator acceptance are assessed using established scales, where these exist, as well as through the evaluation of questionnaire responses and subject comments.

The intention of the experiment is to set up a simulated operations area for the dispatchers to work in. They will be given scenarios in which they are presented with stored "company routes" for a particular city-pair and aircraft type. A diverse set of external weather information sources is represented by a stand-alone display (MOCK), containing the actual historical weather data typically used by dispatchers. There is also the possibility of presenting selected weather data on the route visualization tool.

The company routes have not been modified to avoid the weather except in the case of one additional route generated by the Honeywell prototype flight planning system. The dispatcher will be required to choose the most appropriate and efficient flight plan route in the displayed weather conditions. The route may be modified manually or may be chosen from those automatically displayed.

There will be two flight-planning sessions for each dispatcher within the experiment. One session will utilize the standard AWIN planning tool (hereafter labeled *Concept-A*) and the other will utilize a route visualization (modified AWIN) planning tool (hereafter labeled *Concept-B*). Both tools can display multiple flight plans on the screen. However, the *Concept-A* tool will represent weather in the form of polygon, and the tool will also offer an automatically generated free route that avoids the weather hazards that the dispatchers may choose instead of one of the company routes. The *Concept-B* tool allows the possibility for dispatchers to manually modify an existing route in order to avoid hazards. In both conditions, the dispatchers have access to MOCK containing all the raw weather data that the meteorologists used to generate the polygons for *Concept-A*.

There are two independent variables in this experiment. The first independent variable, *Tool Configuration* has 2 levels, "*Concept-A*" and "*Concept-B*". The second independent variable, *Route* has 12 levels. The 12 *Routes*, or scenarios, were seen by each dispatcher in randomized sequence. Half the dispatchers saw a block of six trials in the *Concept-A* condition followed by six trials in the *Concept-B* condition. The remaining dispatchers saw the *Concept-B* condition first and the *Concept-A* condition second.

The subject pool contained various levels of experience with flight planning tool and meteorology training. Across subjects, there was strong support for the integration of processed weather information in the form of polygons. The experimental results detailed in this report revealed a significant effect of *Concept* in every measurement where the condition was present. The principal benefit of *Concept-A* was the inclusion of weather polygons representing area of hazardous weather. Subjects rated their trust in the polygon definition, boundaries, and severities as high.

Workload of all types was significantly reduced in *Concept-A* over *Concept-B*, where subjects were required to mentally integrate weather and route information across two screens and two applications.

The "Distance Flown in Hazard" metric endeavors to measure dispatchers success level in avoiding hazardous weather, *as defined by the staff meteorologist*. On average, subjects under *Concept-B* flew almost six times as many miles within an area of hazardous weather (as defined by the meteorologist) than subjects under *Concept-A*. Analyzing the data a different way reveals that subjects in *Concept-B* were over six times as likely as those in *Concept-A* to select a route that penetrates areas of weather determined by the meteorologist to be too severe to fly through. It should be noted that there was little difference between *Concept-A* and *Concept-B* in the number of miles once a route that penetrates weather has been chosen (230 miles versus 256 miles, respectively).

On average, across all 12 routes, fuel use in *Concept-B* was 9.0% higher than fuel use under *Concept-A*. Planning time averaged 250 seconds (4:10 min:sec) under *Concept-A*, while *Concept-B* trials averaged 430 seconds (7:10 min:sec).

Subjective comments via questionnaires and Likert scales were another rich source of data. Comparison between *Concept* features were ranked. Issues of trust were also explored. Feedback on the validity of experimental scenarios and data was gathered. Subject comments demonstrated that they clearly felt that the experiment was valid and realistic. The scenarios' validity rested primarily on the fact that real data was used, and the routes were consistent with real operations. The weather came from actual recorded weather, and the company routes were selected from a database of actual routes flown by airlines. Subjects expressed concern that MOCK data, while rated good enough, was missing data they wanted, or at least data in the format they wanted (*e.g.,* AIRMETs).

Subjects also provided formative feedback in the form of comments and answers to questionnaires.

2. Introduction

This document describes the results and analysis of the formal evaluation plan for the Honeywell software tool developed under the NASA AWIN "Weather Avoidance using Route Optimization as a Decision Aid" project. The software tool aims to provide airline dispatchers with a decision aid for selecting optimal routes that avoid weather and other hazards. This evaluation compares and contrasts route selection performance with the AWIN tool to that of subjects using a more traditional dispatcher environment. The evaluation assesses gains in safety, in fuel efficiency of planned routes, and in time efficiency in the pre-flight dispatch process through the use of the AWIN decision aid. In addition, we are interested in how this AWIN tool affects constructs that can be related to performance. The construct of Situation Awareness (SA), workload, trust in an information system, and operator acceptance are assessed using established scales, where these exist, as well as through the evaluation of questionnaire responses and subject comments.

Chapter 3 describes the final experimental procedure used in the formal evaluation. A detailed evaluation plan can be found in [6], which was the basis for the dry-run experiment described in 3.1. Based on the results of the dry-run, the experimental procedure was modified in order to meet experimental objectives, as detailed in section 3.2. The final experiment plan is described in section 3.3

Chapter 4 describes the analysis methodology. Section 4.1 lays out the constraints under which the experiment was designed, often resulting in competing goals. The data collection strategy is detailed in section 4.2. The resulting analysis methodology is described in section 4.3. A derivation of the experimental analysis can be found in Appendix B.

The experimental results are described in Chapter 5. The subject population is described in (section 5.2). Descriptive and empirical results are given for workload (section 5.2), distance flown in hazard (section 5.3), fuel use (section 5.4), planning time (section 5.5), weather sources accessed (section 5.6), situation awareness probes (section 5.7), and subjective questionnaires (section 5.8). Examples of the experimental protocol and experimental briefing materials used can be found in Appendix C and Appendix D, respectively.

Chapter 6 discusses two of the 12 routes in detail to elucidate different routing and weather avoidance strategies in the context of a discussion of what was learned in the experiment.

3. Final Experimental Procedure

This chapter describes the final experimental procedure used in the formal evaluation. The basis of the dry run was a detailed evaluation plan, described in [6], which was developed with input from researchers at Embry-Riddle Aeronautical University. Based on the results of the dry-run, the experimental procedure was modified in order to meet experimental objectives.

3.1 *Dry Run: Lessons Learned*

A dry run was conducted at Embry-Riddle Aeronautical University with four subjects on 15-16 April 2002. The principal lessons learned from the dry-run are the following:

1. Experiment is too lengthy.
2. Weather imagery (MOCK) is sufficient and complete. Subjects did not feel that they were missing any weather information (except PIREPs and TAFs, which were not ready in time)
3. We were unable to modify the commercial flight planner to accept pre-recorded data.

3.1.1 Experimental Length

The original plan called for the entire experiment to last under three and half hours. One principal reason for the dry run was to assess the expected running time. The dry run had four student dispatchers run through all the training and questionnaires, and one trial in each condition. The training took, on average, about as much time as was predicted, however the trials took considerably longer than predicted. It was estimated that a single trial (including post-scenario questionnaire and TLX survey) would last 6 minutes. The AWIN trial lasted 31 minutes and the non-AWIN trial lasted 32 minutes. The other significant underestimate was the time taken to fill out the post-experiment questionnaire (20 minutes predicted vs. 57 minutes actual).

In order to get as much feedback during the dry run as possible, subjects were asked to write down comments to each question they were asked, including the TLX Survey. This also contributed to a longer than expected time to the trials. During the experiment itself, however, subjects were not asked for comments and filled out the TLX survey in less than 30 seconds on average.

3.1.2 MOCK Images

Subjects were repeatedly asked if they could not find any weather information they needed in order to complete their tasks. The only weather data they felt was missing was text weather data such as METARs and TAFs. Text weather data was not ready by the time of the dry run, but was included in the final experiment.

3.1.3 Inability to Import Wind Data

Subjects are asked to select/modify a route to make it as fuel-efficient as possible while avoiding hazardous weather. The experimental plan called for fuel-burn comparisons to be done across conditions, using a single performance model with real wind data to make comparisons. The AWIN tool optimizes a route for fuel-efficiency using real wind data. The candidate commercial flight planner tool can only use a constant (heading and speed) wind field. Thus subjects in the Control case using the commercial flight planner would be trying to optimize for a wind field that is different from the type of wind data used in the evaluation of fuel efficiency in the post-experiment analysis. Clearly, this is a case of evaluating their performance with different criteria than was used in giving them feedback as they performed their task – an unfair and invalid analysis. We could have used a constant wind field within AWIN to make the comparisons valid, but this would have removed a major metric of the experimental objectives – optimizing for fuel-efficiency. The experiment would be reduced to a comparisons in time efficiency (of route modification capabilities) alone.

3.2 *Revisions to the Experiment Plan*

What follows is the two-pronged approach used to revise the experiment in light of what was learned in the dry-run.

3.2.1 The Control Condition

Limitations in the ability of the commercial flight planner to import pre-recorded wind data precluded its use for the control condition. The needed features of the control (non-AWIN) condition are as follows:

- Visualize routes on a US map
- Calculate fuel-burn using actual wind and temp data
- Manually modify routes
- What-if capabilities (change routes, get fuel feedback)
- Lateral and vertical view of routes

The AWIN tool was modified to fulfill the above criteria. Thus the experiment used a modified version of the AWIN tool as the control condition. The principal advantage pertains to the issue of (pre-recorded) real wind data during the trials. The modified version allowed users to manually modify routes, but did not contain the preprocessed polygon weather hazard information created by Embry-Riddle Aeronautical University (ERAU) meteorologists. Both versions did have wind and temperature information. Using a modified version of the AWIN tool in the control condition also reduced training requirements, thus resulting in time savings.

3.2.2 Trial Protocol

The principal factor in the time problem mentioned in Section 3.1.1 is the length of a single trial. The trial protocol was changed significantly to address this problem. Table 1 lists the dry-run trial setup, and the final trial setup.

Table 1. Dry-Run and Final Trial protocol.

Dry-Run	Time (min)	Final	Time (min)
16 trials	-	12 trials	-
Unique Weather Case/trial	-	Unique Wx Case/3 trials	-
		Weather SA: 5 min/3 trials	30
Routing 20min/trial	320	Routing: 7 min / trial	84
Questionnaire: 6 min/trial	96	Questionnaire: 6 min/block	12
TLX: 5 min -> 1 min/trial	16	TLX: 1 min/trial	12
TOTAL	**432 (7:12)**		**138 (2:18)**

The principal features of the final trial protocol are as follows:

a.) The number of trials was reduced to 12 from 16.

b.) There is unique Weather for every three trials

i) Subjects are given five minutes to assess weather (via MOCK) and develop global situation awareness (SA). By giving a dedicated time for SA development, subsequent performance measurements are for routing-specific tasks only.

ii) Once the SA session is over, subjects will be presented with a city-pair to route.

iii) Once they have selected a route, they are given a second city-pair in a different region of the country but for the same weather case.

iv) A third city-pair will follow, also in a geographically distinct (from the other two trials) region of the country

c.) The post-scenario questionnaire was moved from after each trial to after each trial block as the questions were experimental-condition-specific, rather than scenario-specific.

d.) Due to experience of previous evaluations, we expected the time spent on TLX Survey to rapidly approach one minute rather than the five we observed (in addition they were asked to write elaborations during the dry run, which was not the case in the actual experiment).

The sum of these changes allowed the entire experiment to be completed in under four hours. The issue that bears the most scrutiny is the proposal to use the same weather case for two trials. It is a requirement that the trials be independent. By giving subjects a dedicated block of time (5 minutes) to assess weather prior to the beginning of the trial, we are separating out the "Assess Weather in General and Build Weather SA" task from the "Assess Weather As It Impacts the Flight Route" task. Thus if the three trial's city-pairs are in geographically distinct regions subjects to different weather phenomena, the trials will still be independent, since global SA-gathering has occurred outside the trial. In addition, by focussing the trial on local route-specific weather assessment, we can better measure any performance gains pre-processed polygon weather on AWIN affords, potentially a question of greater interest. Experimental analysis will include *Weather Case* to test the assumption that *Weather Case* will not be a significant effect.

3.3 *Revised Experimental Plan*

The proposed changes were accepted by all of the participants, and the experimental plan was revised as a result. This section highlights the experimental plan in light of the revisions.

2.1.1 Hypothesis

Null Hypothesis. The use of AWIN will not reduce dispatcher workload, will not improve flight fuel efficiency and will not reduce the planned route penetration of potentially hazardous weather areas.

Primary Hypothesis. The use of AWIN automated route generation will provide more fuel efficient routes with less dispatcher workload and reduce the planned route penetration of potentially hazardous weather areas.

2.1.2 Operational Scenario

The intention of the experiment is to set up a simulated operations area for the dispatchers to work in. They will be given scenarios in which they are presented with stored routes for a particular city-pair and aircraft type. A diverse set of external weather information sources is represented by a stand-alone display (MOCK), containing the weather data typically used by dispatchers. There is also the possibility of presenting selected weather data on the route visualization tool.

Several pre-selected company routes will be available for each city-pair and are representative of routes flown by major airlines between that city-pair. The company routes are obtained from an ERAU database of flight plans filed for the city-pair. The routes will have been pre-ordered for the exercise in 'fuel efficiency order' but they will not have been modified to avoid the weather except in the case of one route generated by the AWIN system. The dispatcher will be required to choose the most appropriate and efficient flight plan route in the displayed weather conditions. The route may be modified manually or may be chosen from those automatically displayed.

There will be two flight-planning sessions for each dispatcher within the experiment. One session will utilize the standard AWIN planning tool (hereafter labeled *Concept-A*) and the other will utilize a route visualization (modified AWIN) planning tool (hereafter labeled *Concept-B*). Both tools can display multiple flight plans on the screen. However, the *Concept-A* will represent weather in the form of polygons (defined by a meteorologist), and the tool will also offer an automatically generated free route that avoids the weather hazards that the dispatchers may choose instead of one of the company routes. The *Concept-B* tool allows the possibility for dispatchers to manually modify an existing route in order to avoid hazards. In both conditions the dispatchers have access to MOCK containing all the raw weather data that the meteorologists used to generate the polygons for *Concept-A*.

The experiments will make use of stored weather. Each experimental scenario (trial) will involve route selection between a unique city-pair. A unique weather case will be repeated for three trials. Additionally, the scenarios will be in randomized sequence. Of the twelve unique trials, a single participant will conduct six trials in the standard AWIN condition, and six trials in the modified AWIN condition.

For every three trials, participants are given five minutes to assess the weather (via MOCK) and to develop situation awareness (SA). By giving a dedicated time for SA development, subsequent performance measurements are for route-specific tasks only. Once the SA session is over, participants will be presented with three city-pairs to route, one at a time. After each trial (city-pair) they will fill out a TLX Workload survey.

2.1.3 Independent Variables

The first independent variable, *Tool Configuration* has 2 levels, "*Concept-A*" and "*Concept-B*". The second independent variable, *Route* has 12 levels. Note that the *Route* independent variable is nested within the *Weather Case* factor (see section 3.3.2).

3.3.1 Order of trials

The experiment will use a design with the scenarios seen by each dispatcher in randomized sequence. Half the dispatchers will see a block of six trials in the *Concept-A* condition followed by six trials in the *Concept-B* condition. The remaining dispatchers will see the *Concept-B* condition first and the *Concept-A* condition second. (see Appendix A).

3.3.2 City-Pairs

Table 2 illustrates the city-pairs used in the experiment. Note that there are four weather cases, each with three unique routes. Across the twelve routes, no city-pairs share the same two airports, and within a weather case the routes are geographically distinct in order to ensure the (experimental) independence of route.

Table 2. City-pair descriptions.

COMET case Date / Time	Route #	City-Pair		Types of weather encountered
Weather Case #1	1	OKC - MIA	Will Rogers World Airport – Miami International Airport	Convection, turbulence, icing
Jun 06 2001	2	PHX - ORD	Phoenix Sky Harbor International Airport – Chicago O'Hare International Airport	Convection, icing
0015 UTC	3	LAS - YVR	McCarran International Airport (Las Vegas) – Vancouver International Airport	Turbulence, icing
Weather Case #2	4	MIA - LGA	Miami International Airport – New York LaGuardia Airport	Convection
Jun 09 2001	5	SFO - YVR	San Francisco International Airport – Vancouver International Airport	Turbulence, icing
0015 UTC	6	TPA - DFW	Tampa International Airport – Dallas/Fort Worth International Airport	Convection, turbulence, icing
Weather Case #3	7	TPA - AUS	Tampa International Airport – Austin-Bergstrom International Airport (Austin, TX)	Convection, icing
Jun 10 2001	8	MSP – EWR	Minneapolis / St. Paul International Airport – Newark International Airport	Convection, icing
1215 UTC	9	YVR - DEN	Vancouver International Airport – Denver International Airport	Turbulence, icing
Weather Case #4	10	YVR - ORD	Vancouver International Airport – Chicago O'Hare International Airport	Convection
Jun 12 2001	11	MIA – IAD	Miami (FL) International Airport – Washington Dulles International Airport	Convection, icing
0015 UTC	12	PHX – YVR	Phoenix Sky Harbor International Airport – Vancouver International Airport	Turbulence, icing

4. Analysis Methodology

This Chapter describes the data collection and analysis methods used in the experiment.. Section 4.1 discusses the constraints under which the experiment was designed, often resulting in competing goals. The data collection strategy is detailed in section 4.2. The resulting analysis methodology is described in section 4.3. A derivation of the experimental analysis can be found in Appendix B.

4.1 *Experiment Design Constraints*

When considering the design of the experiment, several factors came into play:

- Resource limitations: we had a finite number of subjects and a finite running time
- We wanted all subjects to see both *Concept-A* and *Concept-B* to gather comparative qualitative feedback.
- Within-subjects concerns: primarily the learning effect.
- We only had four weather cases for the twelve routes.

We were limited to 32 subjects and a running time of about four hours per test session. Given that we wanted to run 12 scenarios, there was not enough time to do a full 2 (Concept) x 12 scenarios within-subjects design. Hence the resulting design consists of a pair of fractional factorial designs.

The experiment had several aspects to it, including both quantitative assessment (the controlled experiential evaluation) and qualitative assessment (questionnaires, comparative ratings, trust issues, workload assessment, usability feedback). The qualitative assessment required that subjects see both *Concept* conditions (*Concept-A* and *Concept-B*). Classically, in a within-subjects design, a subject would see a set of routes under both *Concept-A* and those same routes under *Concept-B*. However, there are real concerns about the learning effect in such a design. A subject would already be familiar with the weather and route if they saw the same *Wx Case* / *Route* condition a second time, and therefore the results would be skewed by an order effect (*e.g.,* the planning time data from a particular *Concept* would be different if that trial happened before or after seeing those same conditions with the other *Concept*). Thus the experimental design was such that subjects even though subjects saw all routes, they saw only half the routes in *Concept-A* and the other half of the routes in *Concept-B*.

Having only four weather cases, the experimental design required that each weather case has three routes. *Wx Case 1* is always associated with *Routes* 1-3, *Wx Case* 2 is always associated with *Routes* 4-6 and so on. Thus the experiment is said to have *Routes* nested under *Weather Case* (for more information on nested designs, see [14].

4.2 *Data Collection Design and Implications*

The experimental conditions under which the data was collected is summarized in Table 3.

Table 3. The combinations of experimental variables unique subject groups were run under.

C_A												C_B											
W_A			W_B			W_C			W_D			W_A			W_B			W_C			W_D		
R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}	R_{11}	R_{12}	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}	R_{11}	R_{12}
G_1	G_1	G_1	G_1	G_1	G_1													G_1	G_1	G_1	G_1	G_1	G_1
						G_2	G_2	G_2	G_2	G_2	G_2	G_2	G_2	G_2	G_2	G_2	G_2						

Where:

- G_1 are a group 16 subjects
- G_2 are a group 16 subjects
- NOTE: We randomized the order of *Concept*.
- NOTE: We randomized the order of the *Routes* within a *Wx Case*,
- NOTE: We partially randomized the order of *Wx Cases*.

As seen by the above design, *Route* is nested within *Wx Case*. As a result we have the following observations:

- It is impossible to calculate an interaction term between *Wx Case* (W) and *Route* (R).
- The design assumes that the interaction term W x R is not significant.

In addition, we observe:

- Every subject sees all levels of all variables
- BUT a single subject does NOT see all possible combinations of all variables

The following section describes the analyses conducted on the data that factor in the implications discussed in this section.

4.3 *Analysis*

The data for the independent variables listed in section 2.1.3 will undergo multiple analyses. An overview is given in Table 4, and the ANOVA analyses are described in more detail below.

Table 4. Overview of Planned Analyses of the Data.

Analysis	Description	Motivation
Two (2) *Concept* x 2(3*) Wx Case (Routes)* ANOVA	The data will be divided into two groups (data for *Wx Case* 1-2, and data for *Wx Case* 3-4), and a nested 2x2(3) ANOVA will be performed on each group	Dividing the data in this way allows for analysis simultaneously considering all three independent variables, Concept, Route, and Wx Case where the Route factor is nested under the Wx Case factor
Post-hoc Analysis	Tukey Analysis	For significant variables, we can discern which particular treatments contributed to our rejection of the null hypothesis.

The Experimental Design can be re-organized, as shown in Table 5. Due to the constraints listed in section 4.1, it was not possible to do a pure within-subjects design (where every subject sees every combination of variables) and thus the experiment was complicated by the nested effect of *Wx Case* and *Route*. If the data is split into two groups, one group for *Routes* 1-6 and the second group for *Routes* 7-12, then a straightforward analysis is possible for each group of data.

Table 5. Experimental data collected as for each combination of independent variables.

	Concept A				**Concept B**			
Route	**Wx 1**	**Wx 2**	**Wx 3**	**Wx 4**	**Wx 1**	**Wx 2**	**Wx 3**	**Wx 4**
1	G1				G2			
2	G1				G2			
3	G1				G2			
4		G1				G2		
5		G1				G2		
6		G1				G2		
7			G2				G1	
8			G2				G1	
9			G2				G1	
10				G2				G1
11				G2				G1
12				G2				G1

The resulting ANOVA design is a mixed design with *Concept* as a between-subjects variable, and *Route* and *Wx Case* as within-subject variables (having *Route* effects nested under the *Wx Case* variable). The ANOVA table for each 2 x 2(3) design is given in Table 6. The derivation of the analysis can be found in Appendix B.

Table 6. ANOVA table for a 2x2(3) mixed design.

Source of Variability		DOF Formula	DOF
C	*Concept*	c-1	1
W	*Wx Case*	w-1	1
R (W)	*Routes* w/i *Wx Case* (pooled)	w(r-1)	4
Subjects (C)	*Subjects* w/i *Concept* (pooled)	c(n-1)	30
C x W	Interaction	(c-1)(w-1)	1
Subject(C) x W	Interaction	c(n-1)(w-1)	30
R(W) x C	Interaction	w(r-1)(c-1)	4
Subject(C) x R(W)	Error	cw(n-1)(r-1)	120

4.4 *Metrics*

The post-experiment analysis will be conducted for a series of dispatcher job efficiency measures, weather and hazard penetration measures, and aircraft performance efficiency measures.

Dispatcher Job Efficiency Measures. Metrics to assess job efficiency and workload will consist of subjective and objective measures. Objective measures include: (1) overall planning time in each tool condition, and (2) weather sources accessed during route selection process (*i.e.,* in each trial). Subjective measures will include: (1) TLX score, (2) subject assessment of situation awareness of weather information.

Weather and Hazard Penetration Measures. Routes will have a coefficient generated from the distance flown in weather (above threshold) for the route selected by the subject. This coefficient will be used to rank the weather and hazard avoidance of each planned route.

Aircraft Performance Efficiency Measures. The route chosen will be measured in terms of estimated fuel consumption using a common model for fuel calculations.

A pair-wise comparison will then be carried out of the planned route from the *Concept-A* and the *Concept-B* system on the metrics included above.

5. Results

This Chapter describes the experimental results. Significant results and those further described in text are presented in bold text. The participants in this study, the subject population and the meteorologists are described in section 5.1. Descriptive and empirical results are given for workload (section 5.2), distance flown in hazard (section 5.3), fuel use (section 5.4), planning time (section 5.5), weather sources accessed (section 5.6), situation awareness probes (section 5.7), and subjective questionnaires (section 5.8). Examples of the experimental protocol and experimental briefing materials used can be found in Appendix C and Appendix D, respectively.

5.1 *Participant Characteristics*

The experiment utilized 32 subjects. They were predominantly male (31 males, 1 female) from age 19 to 32, with an average age of 22.5 years. They were asked to write their years of education, where 12 equaled a high school diploma, 16 equaled a four-year college degree, and so on. Years of education varied from 12 to 20 years, with an average of 15.4 years. Thus the average subject in this experiment was a college student entering his or her final (Senior) year. Table 7 summarizes the subject demographic data.

Table 7. Subject demographics.

Measure	Min	Max	Median	Mean	Std. Deviation
Age	19	32	21.0	22.5	3.3
Years Education	12	20	15.0	15.4	1.3

Subjects were asked to detail their experience with pre-flight route selection, including a description of any tools they used. Table 8 details groupings of responses. Experience ranged from none (2) to dispatcher students and instructors (8) to subjects with actual dispatch experience (9), typically in the GA domain. Additionally, three subjects had experience observing or assisting commercial dispatchers in airline operations.

Table 8. Groupings of subject responses on experience with pre-flight route selection.

Experience	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
None				x																											x	
Dispatch classes (student/teacher)	x					x							x	x								x		x	x					x		
Flight / Flight planning experience (GA/commercial))		x				x							x					x	x		x		x					x				x
Communication/assist/observe with flight planner			x		x					x																						

Subjects had a wide experience with flight dispatch tools and resources, as detailed in Table 9. More than half (19) of the subjects had some experience with flight planning tools. Weather information types ranged across the generic “weather reports” to specific responses listing the weather product (TAFs, METARs, en-route charts, etc.). Sources of weather information included specialized software to internet sources.

Table 9. Groupings of subject responses on the tools they have used to do pre-flight route selection.

Tools Used	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Flight Planning software	x	x			x		x	x	x			x	x		x			x		x		x	x	x		x	x	x	x			x
Wx reports			x								x	x						x	x			x	x						x	x		
Internet sources												x					x				x	x										
TAF / METARs / Radar / Sat.			x													x																
Enroute charts						x									x																	
Plotter, sectionals			x					x							x						x			x								

Subjects were asked to detail their experience with meteorology, including a description of any meteorological tools they used. Table 10 describes the experience that subjects listed. Experience with meteorology was primarily through classes at ERAU, where 18 subjects listed taking two semesters of meteorology courses, and another subject reported teaching meteorology. One subject had experience in a meteorology job, and another had experience making go / no-go decisions. Only three subjects listed no meteorology experience.

Table 10. Groupings of subject responses to the question of their experience with meteorology.

Experience	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
None				x															x			x										
Classes	x		x				x		x	x		x	x		x	x		x		x	x			x	x	x	x	x		x		
Teaching classes											x																					
Flight experience	x	x				x		x			x									x					x							
Meteorology job Experience					x																											
Making go/no-go decisions						x																										

Table 11 groups subjects' written responses to the question of their experience with meteorological tools. Fourteen subjects listed the Internet (7), the weather channel (1) or specific weather products (9) as sources of weather information. Additionally, 13 subjects listed a range of computer-based planning tools.

Table 11. Groupings of subject responses to the question of the tools they use for meteorology.

Tools	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Internet								x	x			x			x		x				x					x						
Weather Channel								x																								
Specific Wx products											x			x	x	x	x						x						x	x		x
Computer-based planning tools		x						x							x	x	x						x	x	x			x	x	x	x	x

The staff meteorologist who generated the polygons depicting weather is an Assistant Professor in the Applied Aviation Science (Meteorology) department at the Embry-Riddle Aeronautical University, Daytona Beach, Florida. A graduate student who works with him in this area assisted him.

5.2 *Workload*

One way to assess the effectiveness of decision support tools and their related tasks is to consider the perceived level of workload. In order to measure subjective workload, subjects were asked to fill out a NASA Task Load Index (NASA-TLX) at the end of all trials. The instructions for TLX training can be found in Appendix D. Subjects were asked to rate the level of workload they experienced in completing the tasks involved with selecting a fuel-efficient route while avoiding hazardous weather.

Workload was analyzed in several ways. Subjects completed six TLX rating scales, from 0-10. Each scale represents an individual workload descriptor: mental demand, physical demand, temporal demand, performance, effort, and frustration. Scores over the six scales were averaged to construct a seventh ("Average") rating.

5.2.1 Average Workload

All six workload indices were summed to obtain an average workload score. Table 12 summarizes the descriptive statistics for average workload for each *Concept*, as well as an overall average workload score.

Table 12. Summary of descriptive statistics of TLX average workload assessment.

Metric	Concept A		Concept B.		Overall	
Routes	1-6	7-12	1-6	7-12	1-6	7-12
N of cases	96	96	96	96	192	192
Minimum	0.2	0.3	0.5	0.9	0.2	0.7
Maximum	5.1	6.3	7.7	8.7	7.7	8.7
Median	1.3	2.7	4.0	3.6	2.4	3.0
Mean	1.7	2.6	3.9	3.9	**2.8**	**3.2**
Std. Error	0.11	0.16	0.12	0.21	0.14	0.14
Std. Deviation	1.1	1.5	1.9	2.1	1.9	1.9

Thirty-two subjects took the TLX Workload after each of 12 trials, resulting in 384 scores. Separate analyses were done for routes 1-6 and routes 7-12 (see section 4.3 for a description of the analysis). Overall workload averaged a 2.8 for routes 1-6 and 3.2 for routes 7-12, fairly low on the workload scale of 0-10, suggesting that subjects were not overly taxed.

The results of the Analysis of Variance for *Route* 1-6 and *Route* 7-12 are presented in Table 13. The effect of *Weather Case* is not significant, supporting the design assumption that allowed *Route* to be nested within *Weather Case*.

Table 13. Table of Variance for average workload.

Source of Variability	DOF	F-ratio (P) Routes 1-6	F-ratio (P) Routes 7-12
Concept	1	**F = 529.2 (p < 0.001)**	F = 141.0 (p < 0.001)
Weather Case	1	F = 1.71 (p < 0.193)	F = 0.001 (p < 0.977)
Route (Wx Case)	4	**F = 5.48 (p < 0.001)**	**F = 3.25 (p < 0.014)**
Subjects (Concept))	30	**F = 27.6 (p < 0.001)**	**F = 33.0 (p < 0.001)**
Concept x Wx Case	1	F = 0.18 (p < 0.672)	F = 0.16 (p < 0.686)
Subject(Concept) x Wx Case	30	**F = 1.93 (p < 0.007)**	**F = 1.79 (p < 0.015)**
Concept x Route (Wx Case)	4	F = 1.80 (p < 0.133)	F = 1.79 (p < 0.135)
Error	120		

The effect of *Concept* was found to be strongly significant (F = 529.2, p < 0.001 for routes 1-6; and F = 141.0, p < 0.001 for routes 7-12), as illustrated in Figure 1.

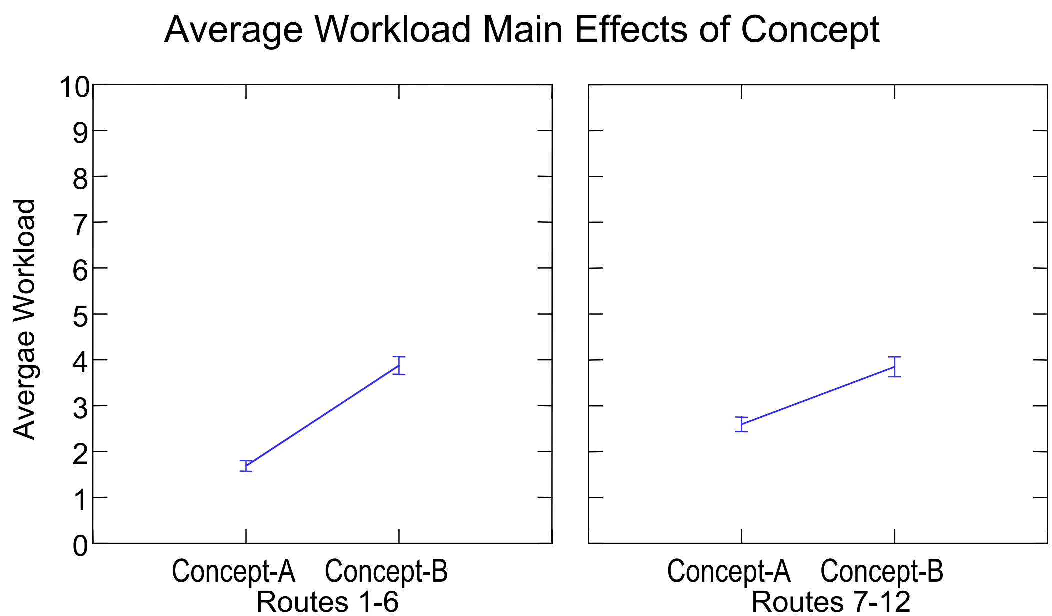


Figure 1. Main effects plot of *Concept* for average workload.

It is not surprising that *Subjects*, which is nested under *Concept*, is then also significant. Since *Subject* is always significant, but only as an artifact of the design of the analysis, it (and its interactions with other independent variables) does not have any explanatory power, and thus will not be discussed further. Figure 2 illustrates the average workload scores for each of 12 *Routes*, under each *Concept*.

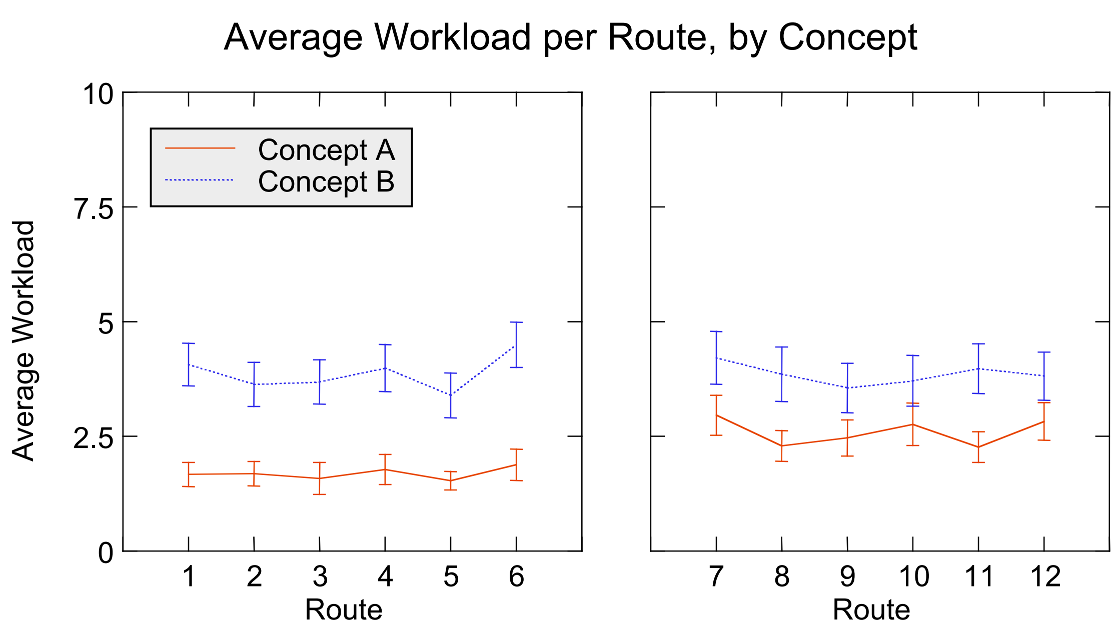


Figure 2. Average workload for 12 routes, broken out by *Concept*.

The results indicate the effect of *Route* is also significant ($F = 5.48$, $p < 0.001$ for routes 1-6; $F = 3.25$, $p < 0.014$ for routes 7-12). Post hoc analysis using Tukey's pairwise comparisons were then performed to find the differences between individual routes. Those comparisons that yielded a probability $p < 0.05$ were considered significant.

For the Tukey comparisons between *Routes 1-6*:

- *Route 6* was different than *Routes 2, 3*, and *5*.

For the Tukey comparisons between *Routes 7-12*:

- *Route 7* is different than *Route 9* and marginally ($p<.066$) from *Route 8*.

Table 14 characterizes the weather encountered on the routes. *Route 6* is similar to Route 1 and 4, in that they all have at least two hazards. Route 6 is different than *Route* 2, 3, and 5, all of which have only one hazard above threshold. Likewise, *Route 7* is similar to *Routes 10, 11*, and *12*, in that all of these routes (except *11*) have multiple hazards. *Routes 8* and *9* have a single hazard each, and are different than *Route 7*.

Table 14. Characterization of weather phenomena encountered on each *Route*.

Route #	City-Pair	Types of weather present en-route	Hazards en-route over threshold		
			Convection	Turbulence	Icing
1	OKC - MIA	Convection, turbulence, icing	2	0	0
2	PHX - ORD	Convection, icing	1	0	0
3	LAS - YVR	Turbulence, icing	0	1	0
4	MIA - LGA	Convection	1	1	0
5	SFO - YVR	Turbulence, icing	0	1	0
6	TPA - DFW	Convection, turbulence, icing	3	0	0
7	TPA - AUS	Convection, icing	2	0	0
8	MSP – EWR	Convection, icing	0	1	0
9	YVR -DEN	Turbulence, icing	0	1	0
10	YVR - ORD	Convection	1	2	1
11	MIA – IAD	Convection, icing	1	0	0
12	PHX – YVR	Turbulence, icing	0	1	1

5.2.2 Mental Workload

The Mental Demand index asks subjects to rate how much mental effort is required to perform the task (*e.g.,* thinking, deciding, remembering). Table 15 summarizes the descriptive statistics for mental workload for each *Concept*, as well as an overall mental workload score.

Table 15. Summary of descriptive statistics of TLX mental Workload assessment.

Metric	Concept A		Concept B.		Overall	
Routes	1-6	7-12	1-6	7-12	1-6	7-12
N of cases	96	96	96	96	192	192
Minimum	0.0	0.0	0.5	0.5	0.0	0.0
Maximum	8.5	8.5	10.0	9.5	10.0	9.5
Median	1.5	3.5	5.5	4.5	3.0	4.0
Mean	2.0	3.4	5.0	4.5	**3.5**	**3.9**
Std. Error	0.17	0.22	0.25	0.26	0.19	0.18
Std. Deviation	1.7	2.2	2.4	2.5	2.6	2.4

Thirty-two subjects took the TLX Workload after each of 12 trials, resulting in 384 scores. Separate analyses were done for routes 1-6 and routes 7-12. Mental workload averaged a 3.5 for routes 1-6 and 3.9 for routes 7-12.

The results of Analysis of Variance for *Route* 1-6 and *Route* 7-12 are given in Table 16. The first thing to note is that the effect of *Weather Case* is not significant, thereby supporting the design assumption that allowed *Route* to be nested within *Weather Case*.

Table 16. Table of Variance for mental workload.

Source of Variability	DOF	F-ratio (P) Routes 1-6		F-ratio (P) Routes 7-12	
Concept	1	**F = 331.5**	**(p < 0.001)**	F = 47.4	(p < 0.001)
Weather Case	1	F = 0.112	(p < 0.739)	F = 0.017	(p < 0.897)
Route (Wx Case)	4	**F = 5.80**	**(p < 0.001)**	**F = 4.36**	**(p < 0.003)**
Subjects (Concept))	30	**F = 15.12**	**(p < 0.001)**	**F = 22.3**	**(p < 0.001)**
Concept x Wx Case	1	F = 1.28	(p < 0.2612)	F = 0.61	(p < 0.437)
Subject(Concept) x Wx Case	30	F = 1.38	(p < 0.117)	**F = 1.60**	**(p < 0.041)**
Concept x Route (Wx Case)	4	**F = 2.80**	**(p < 0.029)**	F = 1.79	(p < 0.136)
Error	120				

The effect of *Concept* is strongly significant (F = 331.5, p < 0.001 for routes 1-6; F = 47.4, p < 0.001 for routes 7-12). The mean score for mental workload under *Concept-A* was 2.0 for *Route* 1-6 and 3.4 for *Route* 7-12. Under *Concept-B*, the means are 5.0 and 4.5, respectively. Since subjects under *Concept-B* must mentally integrate weather information on one screen with route information on another, whereas under *Concept-A* the integration is done in the tool, mental workload is increased under *Concept-B*.

The effect of *Route* is also significant (F = 5.80, p < 0.001 for routes 1-6; F = 4.26, p < 0.003 for routes 7-12). Post hoc analysis using Tukey's pairwise comparisons were then performed to find the differences between individual routes. Those comparisons that yielded a probability p < 0.05 were considered significant.

For the Tukey comparisons between routes 1-6:

- Route 5 was different than *Route* 1, 4, and 6.

For the Tukey comparisons between routes 7-12:

- Route 7 is different than Route 8, 9, and marginally (p<.072) Route 12.

In a similar fashion to the discussion of Average Workload, the principal difference in *Route* seems to be a matter of how many hazards subjects must consider when selecting a route. Route 5 (with one hazard) is significantly different than *Route* 1, 4, and 6 (with two, two, and three hazards respectively). Likewise Route 7 (with two hazards) is significantly different than *Route* 8 or 9 (with one hazard each).

Concept x *Route* is significant (F=2.80, p<0.029) for *Routes* 1-6. Figure 3 illustrates the mental workload scores for each of 12 routes, broken out by *Concept*. Of particular interest is the fact that *Route* 1, 4, and 6 have higher mental workload difference between

Concept-A and *Concept-B* (than the other routes). This can be attributed to their being the only routes (in *Routes* 1-6) that have multiple weather hazards present, leading to more mental integration of route and weather information needed for *Concept-B*.

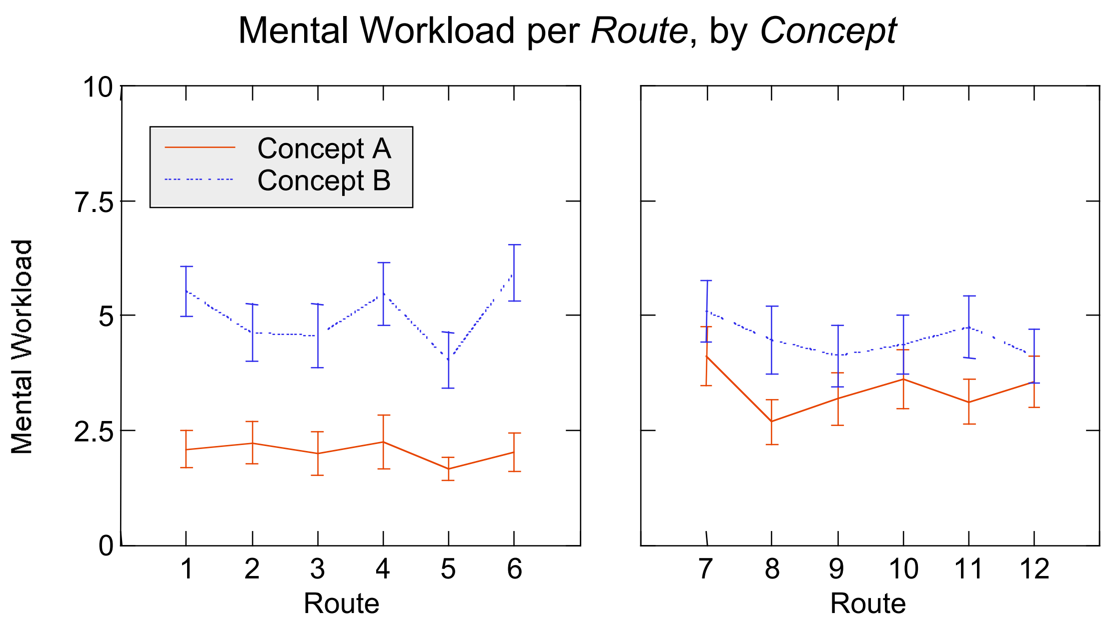


Figure 3. Mental workload for 12 routes, broken out by *Concept*.

5.2.3 Physical Workload

The Physical Demand index asks subjects to rate how much physical effort is required to perform the task (*e.g.,* pushing, pulling, reaching, stretching). Table 17 summarizes the descriptive statistics for physical workload for each *Concept*, as well as an overall physical workload score.

Table 17. Summary of descriptive statistics of TLX physical workload assessment.

Metric	Concept A		Concept B.		Overall	
Routes	1-6	7-12	1-6	7-12	1-6	7-12
N of cases	96	96	96	96	192	192
Minimum	0.0	0.0	0.0	0.0	0.0	0.0
Maximum	6.0	6.0	8.0	10.0	8.0	10.0
Median	0.5	1.0	1.5	1.5	1.0	1.0
Mean	1.0	1.3	2.4	2.7	**1.7**	**2.0**
Std. Error	0.10	0.12	0.21	0.29	0.13	0.17
Std. Deviation	1.0	1.1	2.1	2.9	1.8	2.3

Thirty-two subjects took the TLX Workload after each of 12 trials, resulting in 384 scores. Separate analyses were done for routes 1-6 and routes 7-12. Physical workload averaged a 1.7 for routes 1-6 and 2.0 for routes 7-12. Being a computer-based task, the physical demands of the task were low, and that was reflected in the means reported.

The results of Analysis of Variance for *Route* 1-6 and *Route* 7-12 are given in Table 18. The effect of *Weather Case* is not significant, thereby supporting the design assumption that allowed *Route* to be nested within *Weather Case*.

Table 18. Table of Variance for physical workload.

Source of Variability	DOF	F-ratio (P) Routes 1-6		F-ratio (P) Routes 7-12	
Concept	1	F = 112.7	(p < 0.001)	F = 233.6	(p < 0.001)
Weather Case	1	F = 1.09	(p < 0.299)	F = 0.24	(p < 0.243)
Route (Wx Case)	4	F = 2.53	(p < 0.044)	F = 0.51	(p < 0.726)
Subjects (Concept))	30	F = 15.9	(p < 0.001)	F = 63.5	(p < 0.001)
Concept x Wx Case	1	F = 1.01	(p < 0.318)	F = 0.68	(p < 0.413)
Subject(Concept) x Wx Case	30	F = 1.17	(p < 0.275)	F = 0.91	(p < 0.599)
Concept x Route (Wx Case)	4	F = 1.13	(p < 0.348)	F = 0.69	(p < 0.598)
Error	120				

The effect of *Concept* is strongly significant (F = 112.7, p < 0.001 for routes 1-6; F = 233.6, p < 0.001 for routes 7-12). Figure 4 illustrates the physical workload scores for each of 12 routes, broken out by *Concept*.

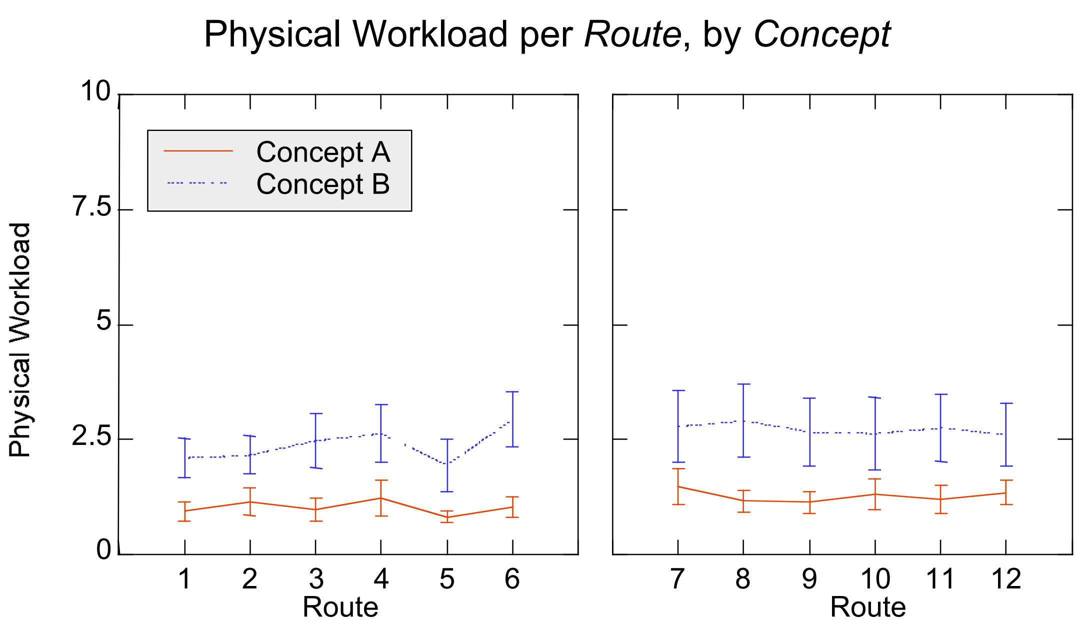


Figure 4. Physical workload for 12 routes, broken out by *Concept*.

The effect of *Route* is also significant for *Route* 1-6 (F = 2.53, p < 0.044) but not for *Route* 7-12. Post hoc analysis using Tukey's pairwise comparisons were then performed to find the differences between individual routes. Those comparisons that yielded a probability p < 0.05 were considered significant. For the Tukey comparisons between routes 1-6 do not show any routes to be significantly different than any other, although Route 5 is marginally different than Route 6 (p < 0.071). Tukey analysis is one of the more conservative post-hoc analysis methods. Route 6 was the only route in the first six to have three hazards, which may account for the increased physical demand of the when compared to Route 5 with only one hazard.

5.2.4 Temporal Demand

The Temporal Demand index asks subjects to rate how much time pressure you feel to complete the task (*e.g.,* relaxed pace or fast and furious?). Table 19 summarizes the descriptive statistics for temporal demand for each *Concept*, as well as an overall temporal demand score.

Table 19. Summary of descriptive statistics of TLX temporal demand assessment.

Metric	Concept A		Concept B.		Overall	
Routes	1-6	7-12	1-6	7-12	1-6	7-12
N of cases	96	96	96	96	192	192
Minimum	0.0	0.0	0.5	0.5	0.0	0.0
Maximum	4.5	8.0	10.0	10.0	10.0	10.0
Median	1.5	3.0	4.5	4.0	2.5	3.5
Mean	1.6	3.2	4.4	4.1	**3.0**	**3.6**
Std. Error	0.12	0.22	0.24	0.27	0.17	0.18
Std. Deviation	1.2	2.2	2.3	2.6	2.3	2.5

Thirty-two subjects took the TLX Workload after each of 12 trials, resulting in 384 scores. Separate analyses were done for routes 1-6 and routes 7-12. Temporal demand averaged a 3.0 for routes 1-6 and 3.6 for routes 7-12.

The results of Analysis of Variance for *Route* 1-6 and *Route* 7-12 are given in Table 20. The effect of *Weather Case* is not significant, thereby supporting the design assumption that allowed *Route* to be nested within *Weather Case*.

Table 20. Table of Variance for temporal demand.

Source of Variability	DOF	F-ratio (P) Routes 1-6		F-ratio (P) Routes 7-12	
Concept	1	**F = 375.0**	**(p < 0.001)**	F = 32.4	(p < 0.001)
Weather Case	1	F = 0.07	(p < 0.786)	F = 1.20	(p < 0.275)
Route (Wx Case)	4	**F = 3.05**	**(p < 0.020)**	**F = 3.66**	**(p < 0.008)**
Subjects (Concept))	30	**F = 15.0**	**(p < 0.001)**	**F = 20.0**	**(p < 0.001)**
Concept x Wx Case	1	F = 0.118	(p < 0.731)	F = 0.21	(p < 0.649)
Subject(Concept) x Wx Case	30	**F = 2.24**	**(p < 0.001)**	F = 1.51	(p < 0.061)
Concept x Route (Wx Case)	4	**F = 2.88**	**(p < 0.025)**	F = 1.10	(p < 0.362)
Error	120				

The effect of *Concept* is strongly significant (F = 375.0, p < 0.001 for *Routes* 1-6; F = 32.4, p < 0.001 for *Routes* 7-12). It was anticipated that under *Concept-B*, temporal demand would be the significant since subjects must manually modify a company route to avoid weather, and this can be a time consuming process. There was a significant jump in temporal demand for both *Routes* 1-6 and *Routes* 7-12.

The effect of *Route* is also significant (F = 3.05, p <0.020 for routes 1-6; F = 3.66, p < 0.008 for routes 7-12). Post hoc analysis using Tukey's pairwise comparisons were then performed to find the differences between individual routes. Those comparisons that yielded a probability p < 0.05 were considered significant. As with the other indices, the principal difference appear to be the number of hazards en-route.

For the Tukey comparisons between routes 1-6:

- Route 5 was different than Route 6.

For the Tukey comparisons between routes 7-12:

- Route 7 is different than Route 8 and 9.

The *Concept* x *Route* interaction is significant ($F=2.88$, $p<0.025$) for *Routes* 1-6. Of particular interest is the fact that Route 1 and 6 had higher temporal demand under *Concept-B* than *Concept-A* as compared to *Routes* 2, 3, 4, and 5. Presumably, this is because they are the only two routes with all three types of weather hazards represented, and therefore they demanded more consideration (time) from subjects when constructing new routes. Figure 5 illustrates the average temporal demand scores for each of 12 routes, broken out by *Concept*.

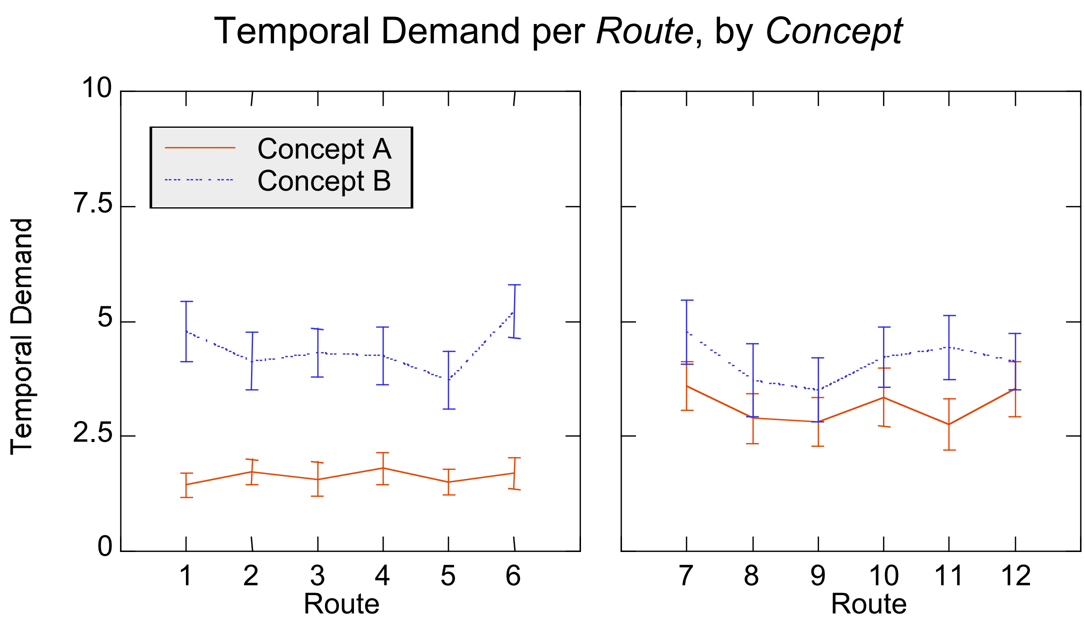


Figure 5. Temporal demand for 12 routes, broken out by *Concept*.

5.2.5 Performance

The Performance index asks subjects to rate how successful they were in completing the task. Table 21 summarizes the descriptive statistics for performance for each *Concept*, as well as an overall performance score.

Table 21. Summary of descriptive statistics of TLX performance assessment.

Metric	Concept A		Concept B.		Overall	
Routes	1-6	7-12	1-6	7-12	1-6	7-12
N of cases	96	96	96	96	192	192
Minimum	0.0	0.0	0.5	1	0.0	0.0
Maximum	7.5	7.0	8.5	8.5	8.5	8.5
Median	1.5	1.5	2.5	3.5	2.0	2.5
Mean	1.8	2.1	2.8	3.7	2.3	**2.9**
Std. Error	0.17	0.18	0.18	0.19	0.13	0.14
Std. Deviation	1.6	1.8	1.7	1.8	1.8	2.0

Thirty-two subjects took the TLX Workload after each of 12 trials, resulting in 384 scores. Separate analyses were done for routes 1-6 and routes 7-12. Performance averaged a 2.3 for routes 1-6 and 2.9 for routes 7-12.

The results of Analysis of Variance for *Route* 1-6 and *Route* 7-12 are given in Table 22. As also shown in the temporal demand analysis, the effect of *Weather Case* is not significant, thereby supporting the design assumption that allowed *Route* to be nested within *Weather Case*.

Table 22. Table of Variance for performance.

Source of Variability	DOF	F-ratio (P) Routes 1-6		F-ratio (P) Routes 7-12	
Concept	1	**F = 45.9**	**(p < 0.001)**	**F = 150.2**	**(p < 0.001)**
Weather Case	1	F = 0.45	(p < 0.506)	F = 0.91	(p < 0.341)
Route (Wx Case)	4	**F = 2.96**	**(p < 0.022)**	**F = 4.00**	**(p < 0.004)**
Subjects (Concept))	30	**F = 9.83**	**(p < 0.001)**	**F = 21.5**	**(p < 0.001)**
Concept x Wx Case	1	F = 3.726	(p < 0.056)	F = 0.50	(p < 0.481)
Subject(Concept) x Wx Case	30	**F = 2.45**	**(p < 0.001)**	F = 1.28	(p < 0.173)
Concept x Route (Wx Case)	4	F = 0.77	(p < 0.550)	F = 0.46	(p < 0.767)
Error	120				

The effect of *Concept* is strongly significant (F = 45.9, p < 0.001 for routes 1-6; F = 150.2, p < 0.001 for routes 7-12). The Performance scale is rated from Good (0) to Poor (10), so a lower score means a better rating of performance. Thus, subjects felt they performed the task better under *Concept-A* (1.8 for *Route* 1-6, and 2.1 for *Route* 7-12) versus *Concept-B* (2.8 for *Route* 1-6, 3.7 for *Route* 7-12). Figure 6 illustrates the average perceived performance scores for each of 12 *Routes*, broken out by *Concept*.

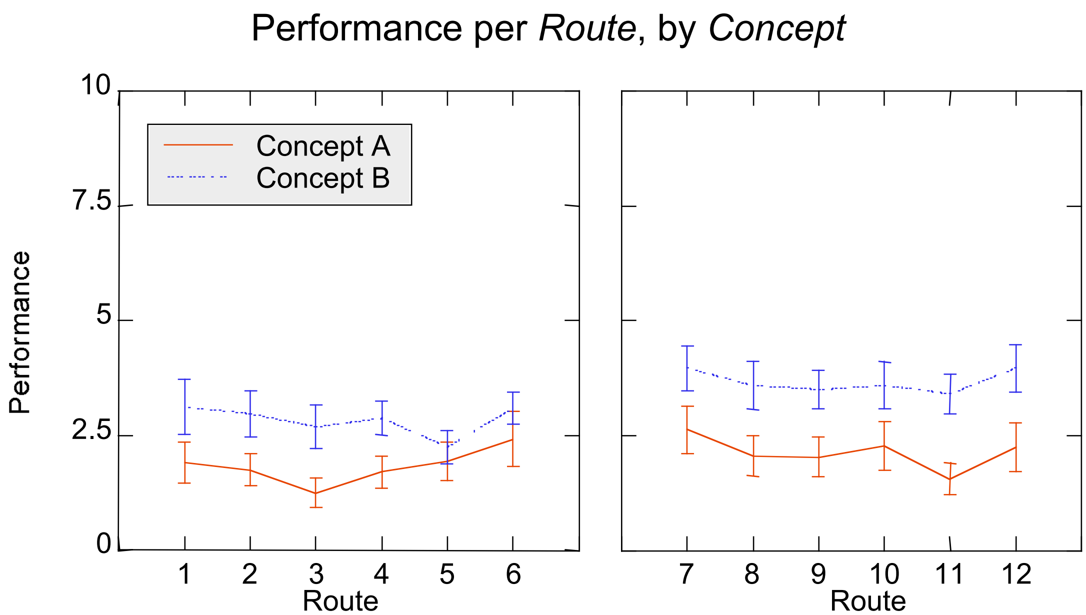


Figure 6. Performance for 12 routes, broken out by *Concept*.

The effect of *Route* is also significant (F = 2.96, $p < 0.022$ for routes 1-6; F = 4.00, $p < 0.004$ for routes 7-12). Post hoc analysis using Tukey's pairwise comparisons were then performed to find the differences between individual routes. Those comparisons that yielded a probability $p < 0.05$ were considered significant. As with the other indices mentioned previously, the difference appears to be the number of hazards en-route.

For the Tukey comparisons between routes 1-6:

- Route 3 was different than Route 6.

For the Tukey comparisons between routes 7-12:

- Route 11 is different than Route 7 and marginally ($p < 0.053$) from Route 12.

5.2.6 Effort

The Effort index asked each subject to rate how hard they worked to complete the task. Table 23 summarizes the descriptive statistics for effort for each *Concept*, as well as an overall effort score for *Route* 1-6 and 7-12.

Table 23. Summary of descriptive statistics of TLX effort assessment.

Metric	Concept A		Concept B.		Overall	
Routes	1-6	7-12	1-6	7-12	1-6	7-12
N of cases	96	96	96	96	192	192
Minimum	0.5	0.0	0.5	0.5	0.5	0.0
Maximum	9.5	9.0	10.0	10.0	10.0	10.0
Median	1.5	3.0	5.5	4.5	3.0	3.5
Mean	2.4	3.2	5.1	4.7	**3.7**	**4.0**
Std. Error	0.23	0.25	0.27	0.27	0.20	0.19
Std. Deviation	2.2	2.4	2.7	2.6	2.8	2.6

Thirty-two subjects took the TLX Workload after each of 12 trials, resulting in 384 scores. Separate analyses were done for routes 1-6 and routes 7-12. Effort averaged a 3.7 for routes 1-6 and 4.0 for routes 7-12.

The results of Analysis of Variance for *Route* 1-6 and *Route* 7-12 are given in Table 24. Again, the effect of *Weather Case* is not significant, thereby supporting the design assumption that allowed *Route* to be nested within *Weather Case*.

Table 24. Table of Variance for effort.

Source of Variability	DOF	F-ratio (P) *Route* 1-6		F-ratio (P) *Route* 7-12	
Concept	1	**F = 236.1**	**(p < 0.001)**	**F = 70.1**	**(p < 0.001)**
Weather Case	1	F = 1.87	(p < 0.174)	F = 0.00	(p < 0.977)
Route (Wx Case)	4	F = 1.63	(p < 0.171)	F = 0.50	(p < 0.738)
Subjects (Concept))	30	**F = 20.4**	**(p < 0.001)**	**F =20.4**	**(p < 0.001)**
Concept x Wx Case	1	F = 0.96	(p < 0.329)	F = 2.04	(p < 0.155)
Subject(Concept) x Wx Case	30	F = 1.14	(p < 0.308)	F = 1.31	(p < 0.153)
Concept x Route (Wx Case)	4	F = 0.33	(p < 0.860)	F = 2.33	(p < 0.060)
Error	120				

The effect of *Concept* is strongly significant (F = 236.1, p < 0.001 for routes 1-6; F = 70.1, p < 0.001 for routes 7-12). Of all the workload measures, it was expected that under *Concept-B* effort would be the greatest, since subjects must not only mentally integrate mentally the weather and route, but also manually modify a company route to avoid weather. This would increase both the temporal demand of the task (discussed earlier) and the effort required to successfully complete it. Indeed for *Concept-B* effort was high when compared to the other indices (5.1 for *Routes* 1-6, 4.7 for *Routes* 7-12), where effort for *Concept-A* was significantly lower (2.4 for *Routes* 1-6, 3.2 for *Routes* 7-12). Figure 7 illustrates the average effort scores for each of 12 *Routes*, broken out by *Concept*.

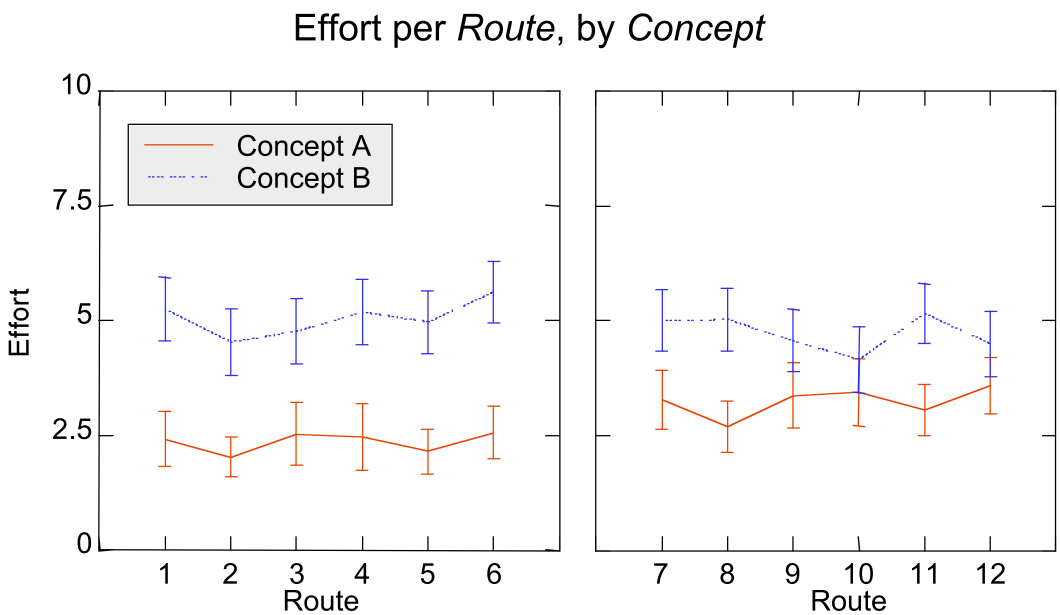


Figure 7. Effort for 12 routes, broken out by *Concept*.

The effect of *Route* was not significant for *Route* 1-6 and 7-12.

5.2.7 Frustration

The Frustration index asks subjects to rate how aggravated or annoyed versus secure or content you feel about accomplishing the task. Table 25 summarizes the descriptive statistics for frustration for each *Concept*, as well as an overall frustration score.

Table 25. Summary of descriptive statistics of TLX frustration assessment.

Metric	Concept A		Concept B.		Overall	
Routes	**1-6**	**7-12**	**1-6**	**7-12**	**1-6**	**7-12**
N of cases	96	96	96	96	192	192
Minimum	0.0	0.0	0.0	0.5	0.0	0.0
Maximum	5.0	8.0	8.5	9.5	8.5	9.5
Median	0.5	2.0	3.0	2.5	1.9	2.5
Mean	1.3	2.4	3.6	3.4	**2.4**	**2.9**
Std. Error	0.11	0.20	0.25	0.27	0.16	0.17
Std. Deviation	1.1	2.0	2.4	2.7	2.2	2.4

Thirty-two subjects took the TLX Workload after each of 12 trials, resulting in 384 scores. Frustration averaged a 2.4 for routes 1-6 and 2.9 for routes 7-12, indicating that subjects felt relatively secure in accomplishing the task.

The results of Analysis of Variance for *Route* 1-6 and *Route* 7-12 are given in Table 26. Once again, the effect of *Weather Case* is not significant, thereby supporting the design assumption that allowed *Route* to be nested within *Weather Case*.

Table 26. Table of Variance for frustration.

Source of Variability	DOF	F-ratio (P) Routes 1-6		F-ratio (P) Routes 7-12	
Concept	1	**F = 238.1**	**(p < 0.001)**	**F = 38.2**	**(p < 0.001)**
Weather Case	1	F = 1.47	(p < 0.227)	F = 0.02	(p < 0.884)
Route (Wx Case)	4	F =1.44	(p < 0.224)	F = 1.74	(p < 0.146)
Subjects (Concept))	30	**F =15.3**	**(p < 0.001)**	**F = 21.6**	**(p < 0.001)**
Concept x Wx Case	1	F = 0.39	(p < 0.534)	F = 0.05	(p < 0.833)
Subject(Concept) x Wx Case	30	F = 1.52	(p < 0.058)	F = 2.67	(p < 0.001)
Concept x Route (Wx Case)	4	F = 0.20	(p < 0.936)	F = 0.61	(p < 0.654)
Error	120				

The effect of *Concept* is strongly significant (F = 238.1, p < 0.001 for routes 1-6; F = 38.2, p < 0.001 for routes 7-12). Under *Concept-A*, subjects rated their frustration as 1.3 for *Routes* 1-6 and 2.4 for *Routes* 7-12. The jump can be attributed to the increased complexity of the hazards and routes for *Routes* 7-12 when compared to *Routes* 1-6. Frustration levels increased even more under *Concept-B*, to 3.6 for *Routes* 1-6 and 3.4 for *Routes* 7-12. Figure 8 illustrates the average frustration scores for each of 12 routes, broken out by *Concept*. The effect of *Route* was not significant.

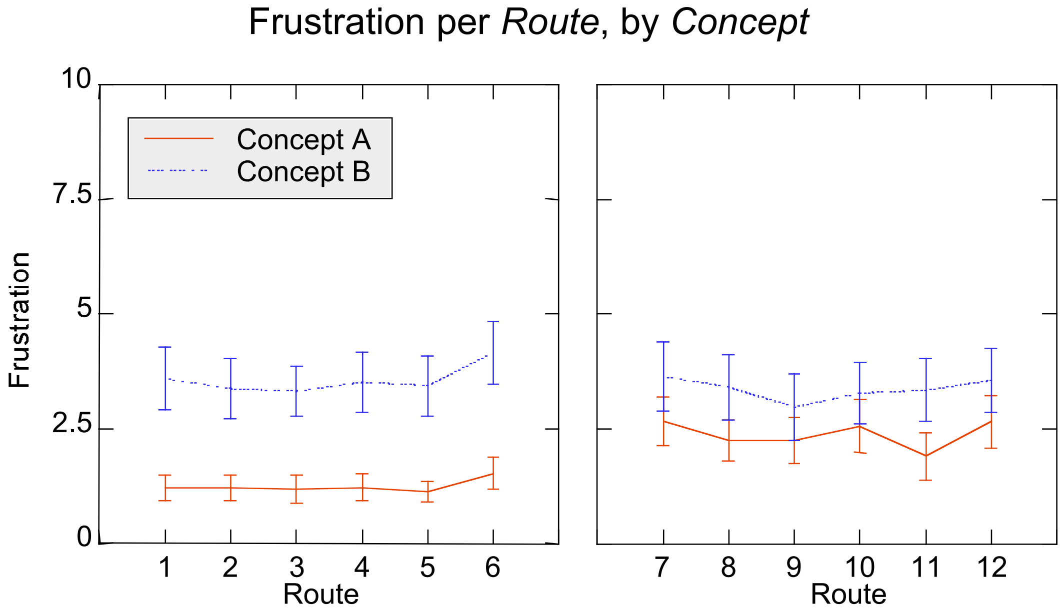


Figure 8. Frustration for 12 routes, broken out by *Concept*.

5.3 *Distance in Hazard*

Dispatchers were told to minimize fuel while choosing a route that avoids hazardous weather. Under *Concept-B,* subjects were given raw weather information in the form of images within MOCK. Under *Concept-A*, in addition to access to MOCK, subjects saw weather in the form of polygons with defined boundaries and severity levels. The polygons were drawn by a meteorologist using the same raw weather information that subjects under both *Concept* conditions had access to during the trials. In both conditions, subjects were trained on the rules the meteorologist used to define hazard severity and boundaries. Thus ideally subjects under both conditions were using the same set of rules when deciding what routes avoided weather deemed too severe to fly through.

The "Distance Flown in Hazard" metric endeavors to measure dispatchers success level in avoiding hazardous weather, *as defined by the staff meteorologist*. By superimposing the polygon hazards on the selected route for a given subject on a given trial, one can calculate the number of miles that are flown inside a weather polygon (above severity threshold). Of course subjects under Concept B did not have access to the polygons when selecting/modifying their route. However, they had access to the raw weather data and were trained on the rules for determining weather severity and boundaries. Table 27 contains the minimum, maximum, mean, and standard deviation of the "distance flown in hazard" per *Route* per *Concept*, as well as aggregates for route 1-6, 7-12, and 1-12

Table 27. Summary of descriptive statistics of "distance flow in hazard" under *Concept-A* and *Concept-B* for each *Route*.

Route	N		Minimum		Maximum		Mean			Std. Dev.		
Concept	**A**	**B**	**A**	**B**	**A**	**B**	**A**	**B**	**% diff**	**A**	**B**	**% diff**
1	16	16	0	0	241	177	26.1	99.5	**280.7**	45.3	56.2	**24.1**
2	16	16	0	0	222	222	13.9	61.4	**342.4**	-	69.6	**-**
3	16	16	0	0	311	356	19.4	286.2	**1372**	-	13.7	**-**
4	16	16	0	0	196	422	24.5	202.7	**727.1**	0.0	98.1	**-**
5	16	16	0	0	326	348	20.4	245.3	**1104**	-	7.3	**-**
6	16	16	0	0	176	221	44.0	123.1	**179.8**	0.0	62.2	
Total 1-6	**96**	**96**	**0**	**0**	**326**	**422**	**25**	**170**	**586.5**	**71**	**133**	**87.3**
7	16	16	0	0	357	622	29.9	287.2	**859.2**	166.2	182.7	**9.9**
8	16	16	0	0	0	172	0.0	101.4	**-**	-	4.4	**-**
9	16	16	0	0	308	398	19.3	221.4	**1050**	-	26.8	**-**
10	16	16	0	0	424	544	64.7	220.0	**239.8**	68.5	87.0	**27**
11	16	16	0	0	9	154	1.7	23.6	**1297**	0.0	63.3	**-**
12	16	16	0	0	416	626	52.0	352.0	**576.9**	0.0	70.8	**-**
Total 7-12	**96**	**96**	**0**	**0**	**424**	**626**	**28**	**201**	**621.5**	**96**	**188**	**95.8**

On average, subjects under *Concept-B* flew almost six times as many miles within an area of hazardous weather (as defined by the meteorologist) than subjects under *Concept-A*. For *Route* 1-6, subjects in *Concept-A* averages 25 miles flown within a hazard, where the subjects under *Concept-B* averaged 170 miles (a 587% increase). Similarly, for *Route* 7-12, subjects in *Concept-A* averages 28 miles flown within a hazard, where the subjects under *Concept-B* averaged 201 miles (a 622% increase). Standard deviation also showed a difference between *Concept-A* and *Concept-B* of 87% for *Route* 1-6 and 95.8% for *Route* 7-12.

The results of Analysis of Variance for *Route* 1-6 and *Route* 7-12 are given in Table 28.

Table 28. Table of Variance for "distance flown in hazard."

Source of Variability	DOF	F-ratio (P) Routes 1-6		F-ratio (P) Routes 7-12	
Concept	1	**F = 151.1**	**(p < 0.001)**	**F = 127.6**	**(p < 0.001)**
Weather Case	1	**F = 4.70**	**(p < 0.032)**	F = 0.356	(p < 0.552)
Route (Wx Case)	4	**F =10.3**	**(p < 0.001)**	**F = 17.6**	**(p < 0.001)**
Subjects (Concept))	30	**F =2.35**	**(p < 0.001)**	**F = 3.78**	**(p < 0.001)**
Concept x Wx Case	1	F = 1.78	(p < 0.184)	F = 0.83	(p < 0.364)
Subject(Concept) x Wx Case	30	F = 1.24	(p < 0.208)	F = 1.16	(p < 0.279)
Concept x Route (Wx Case)	4	**F = 11.9**	**(p < 0.001)**	**F = 9.09**	**(p < 0.001)**
Error	120				

The effect of *Weather Case* is significant for *Route* 1 to 6, with subjects flying smaller distance in hazards for Weather *Case* A (an average of 84 miles) than for *Weather Case* B (an average of 110 miles). This is attributable to *Weather Case A* having fewer en-route hazards over threshold (four) than *Weather Case B* (six).

The effect of *Concept* is strongly significant (F = 151.1, p < 0.001 for *Routes* 1-6; F = 127.6, p < 0.001 for *Routes* 7-12), as expected given the averages in Table 27. Figure 9 illustrates the average distance flown in hazard for each of 12 routes, broken out by *Concept*. *Concept x Route* is significant ((F = 11.9, p < 0.001 for *Routes* 1-6; F = 9.09, p < 0.001 for *Routes* 7-12), as distance flown in hazard varies widely by Route for Concept B, but very little for *Concept-A*. This is likely due to the fact that under *Concept-A*, the vast majority of the subjects (170 of 192, see next paragraph) chose a route with zero distance flown in hazard.

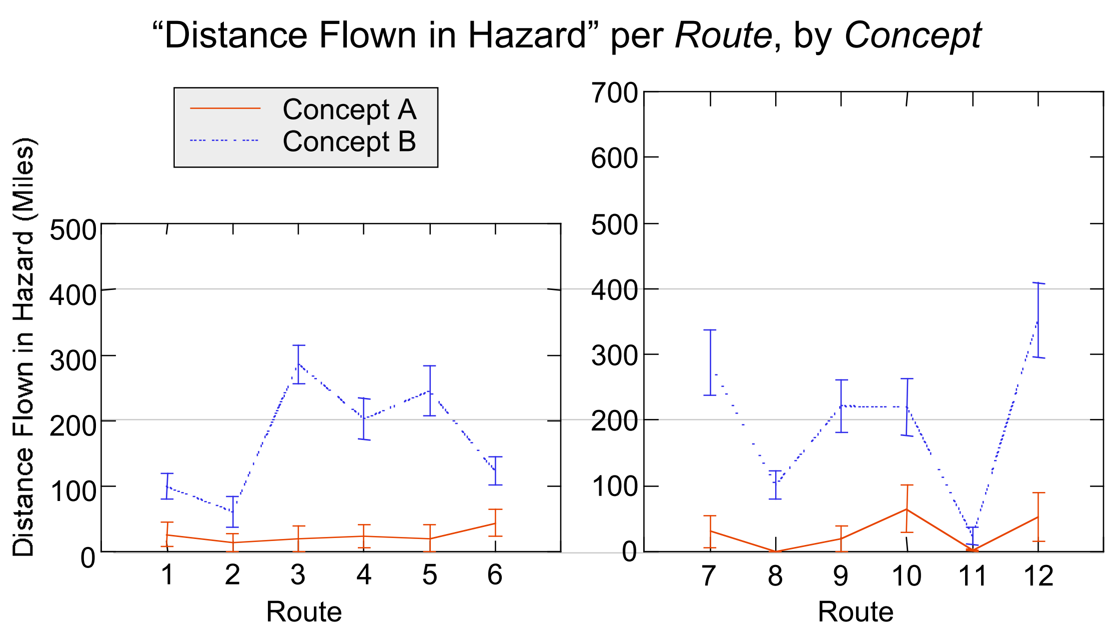


Figure 9. Average Distance flown in hazard, on each route, under Concept A and B.

There are several things of note in Table 27. First of all one might have expected the average distance flown in hazard under *Concept-A* to be zero, since subjects had a route available that was guaranteed to avoid all weather over severity threshold. Clearly, despite its availability, not all subjects in *Concept-A* selected that route, instead selecting routes that penetrated hazardous weather polygons. Conversely, the Table does not tell us how many subjects in *Concept-B* were successful in constructing routes that avoided weather. Table 29 attempts to answer those questions by listing the number of subjects that selected a route that penetrated a hazard above threshold.

Table 29. Characterization of subjects that penetrated hazards.

Route	Concept A				Concept B			
	N	# subjects that penetrated a hazard	% subjects	Average penetration	N	# subjects that penetrated a hazard	% subjects	Average penetration
1	16	2	12.5	209.0	16	12	75.0	132.6
2	16	1	6.3	222.0	16	6	37.5	163.7
3	16	1	6.3	311.0	16	14	87.5	327.1
4	16	2	12.5	196.0	16	14	87.5	231.6
5	16	1	6.3	326.0	16	12	75.0	327.1
6	16	4	25	176.0	16	13	81.3	151.5
7	16	2	12.5	239.5	16	15	93.8	306.3
8	16	0	0	0.0	16	10	62.5	162.3
9	16	1	6.3	308	16	11	68.8	322.0
10	16	3	18.8	345.3	16	11	68.8	320.0
11	16	3	18.8	9.0	16	9	56.3	41.9
12	16	2	12.5	416	16	12	75.0	469.3
Total	**192**	**22**	**11.5**	**229.8**	**192**	**139**	**72.4**	**256.0**

Of the 192 trials under *Concept-A*, a route that penetrated a hazard was selected only 22 times, or approximately 11.5% of the time. This is somewhat surprising because, as stated before, subjects under *Concept-A* had access to a route pre-calculated to avoid hazardous weather. Subjects using the tool under the *Concept-B* condition penetrated a hazard in 139 of 192 trials (74.2% of the time). In other words, subjects were over six times as likely to select a route that penetrates areas of weather determined by the meteorologist to be too severe to fly through. It should be noted that there was little difference between *Concept-A* and *Concept-B* in the number of miles once a route that penetrates weather has been chosen (230 miles versus 256 miles, respectively).

Subjects in the *Concept-B* conditions were the most successful in avoiding hazards for *Route* 2 (10 of 16) and the least successful in *Route* 7 (1 of 16). The two routes are illustrated in Figure 10. Route 2 only features one hazard (convection) above threshold along the company routes, and it is relatively small in size. Route 7, however, has four large areas of hazard over threshold (although the two icing polygons have tops low enough that they do not intersect the company routes).

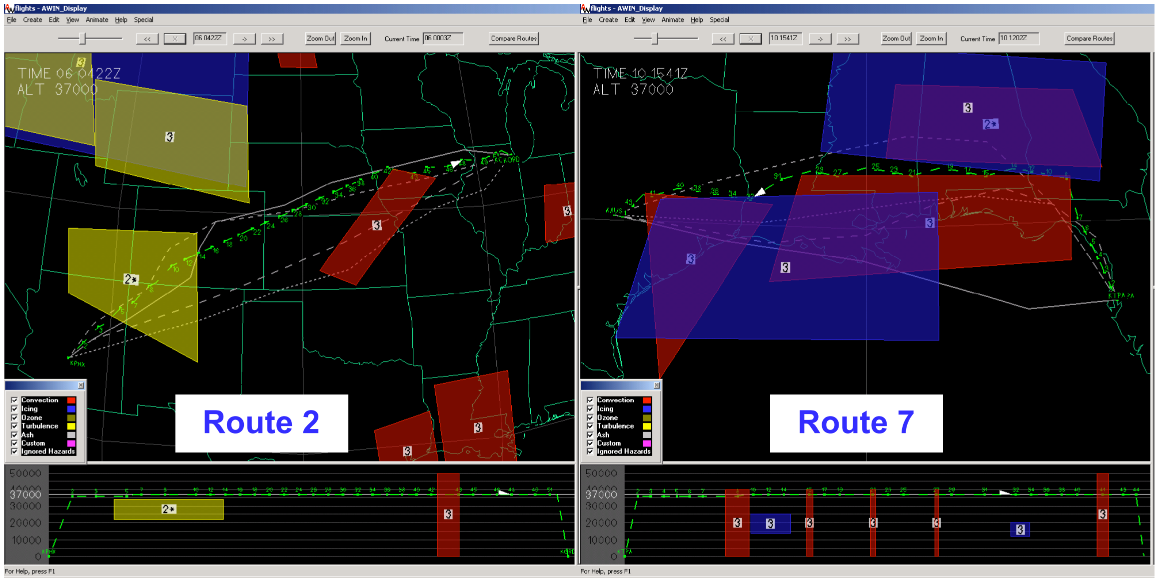


Figure 10. *Route* 2 and *Route* 7.

The next logical question to ask is what kind of hazards are being penetrated, how many hazards are being penetrated, and what percentage of possible hazards are being penetrated. Table 30 answers the first question by listing the total number of hazards (of each type) that were penetrated by subjects as they conducted the experiment. Under *Concept-A*, subjects penetrated a total of 28 hazards (13 Convection, 10 turbulence, and 5 icing). Subjects in *Concept-B* penetrated 167 hazards (65 convection, 82 turbulence, and 20 icing).

Table 30. Number of Hazards penetrated per route for all subjects.

Route	Number of hazards penetrated							
	Concept A				Concept B			
	Conv.	Turb.	Icing	Total	Conv.	Turb.	Icing	Total
1	2			2	12			12
2	1			1	6			6
3		1		1		14		14
4		2		2	5	13		18
5		1		1		12		12
6	4			4	14			14
7	2			2	17			17
8				0		10		10
9		1		1		10		10
10	1	3	3	7	2	11	8	21
11	3			3	9			9
12		2	2	4		12	12	24
Total	**13**	**10**	**5**	**28**	**65**	**82**	**20**	**167**

In order to answer the question of what percentage of hazards of each type are being penetrated, we must first total up how many hazards of each type there are. Table 31 lists the total number of hazards that impinge on the company routes of each of the *Routes* 1-12. It also breaks how many of each type of hazards are in each *Route*. At the bottom it totals the total number of hazards of each type a subject will see during the course of conducting 12 trials corresponding to *Route* 1-12. Finally the last row lists the number of hazards a set of 16 subjects will see (and therefore the number of subjects who saw each route in one Concept condition). Thus, for example, there are 336 possible hazards (176 convection, 128 turbulence, and 32 icing) across the 16 subjects under *Concept-A* for *Route* 1-12.

Table 31. Number of hazards of each type in a route. The last line is the total number of hazards that the 16 subjects saw under one *Concept* condition.

Route	# hazards	Convection	Turbulence	Icing
1	2	2	0	
2	1	1	0	
3	1	0	1	
4	2	1	1	
5	1	0	1	
6	3	3	0	
7	2	2	0	
8	1	0	1	
9	1	0	1	
10	4	1	2	1
11	1	1	0	
12	2	0	1	1
Total per subject	**21**	**11**	**8**	**2**
Total for all 16 subjects (per Concept)	**336**	**176**	**128**	**32**

Finally, we can calculate the percentage of hazards of each type that were penetrated under both *Concept* conditions. Table 32 summarizes the results. On average, under *Concept-A*, subjects penetrated 8.3% of the possible hazards. More specifically, under *Concept-A* subjects penetrated 7.4% of convection hazards, 7.8% of turbulence hazards, and 15.6% icing hazards. In contrast, on average under *Concept-B*, subjects penetrated 49.7 % of the possible hazards. More specifically, under *Concept-B* subjects penetrated 36.9% of convection hazards, 64.1% of turbulence hazards, and 62.5% icing hazards. The higher incidence of turbulence and icing hazard penetration is due to subjects preferred strategy of lateral deviation over vertical route modification. In most cases the turbulence and icing hazards in *Route* 1-12 could be avoided with a change (upwards) of altitude, but few subjects utilized this strategy, preferring instead to modify the route laterally.

Table 32. Percentage of the possible hazards that were penetrated, for all subjects.

	Concept A				Concept B			
	Conv.	**Turb.**	**Icing**	**Total**	**Conv.**	**Turb.**	**Icing**	**Total**
# penetrated	13	10	5	28	65	82	20	167
Total possible (x16)	176	128	32	336	176	128	32	336
% penetrated	**7.4%**	**7.8%**	**15.6%**	**8.3%**	**36.9%**	**64.1%**	**62.5%**	**49.7%**

5.4 *Fuel Use*

The experiment had each subject dispatch 12 distinct routes. The routes varied in length as each featured a different, unique city pair. Thus the fuel numbers will be highly dependent on each city pair (*i.e., Route*). Table 33 contains the minimum, maximum, mean, and standard deviation of the fuel use per route, as well as aggregates for route 1-6, 7-12, and 1-12.

Table 33. Summary of descriptive statistics of fuel use under *Concept-A* and *B* for each *Route*.

Route	N		Minimum Fuel		Maximum Fuel		Mean			Std. Dev.		
Concept	A	B	A	B	A	B	A	B	% diff	A	B	% diff
1	16	16	13455	14229	15383	16239	13686	15271	**11.6**	637	419	**-34.2**
2	16	16	15204	15173	17119	17563	15514	16502	**6.4**	687	668	**-2.8**
3	16	16	13082	13245	14675	15350	13182	14554	**10.4**	398	566	**42.2**
4	16	16	12219	13092	14695	16740	12529	14445	**15.3**	846	968	**14.4**
5	16	16	9794	9803	10567	11073	9842	10414	**5.8**	193	365	**89.1**
6	16	16	12002	11923	12397	15766	12127	13100	**8.0**	111	1142	**928.1**
Total 1-6	**96**	**96**	**9794**	**9803**	**17119**	**17563**	**12813**	**14047**	**9.6**	**1801**	**2058**	**14.3**
7	16	16	11899	11899	13293	16285	12757	13179	**3.3**	262	1137	**334.0**
8	16	16	10608	10633	10608	13833	10608	11700	**10.3**	0	800	**inf**
9	16	16	12753	12738	13962	19525	12753	14393	**12.9**	302	1688	**458.9**
10	16	16	18403	18384	21738	21628	18863	20157	**6.9**	1028	789	**-23.2**
11	16	16	10717	10659	11467	17942	10858	11962	**10.2**	302	1661	**450.0**
12	16	16	15226	15581	16426	17996	15373	16667	**8.4**	411	621	**51.1**
Total 7-12	**96**	**96**	**10608**	**10633**	**21738**	**21628**	**13547**	**14676**	**8.3**	**2901**	**3194**	**10.1**
Total 1-12	**192**	**192**	**9794**	**9803**	**21738**	**21628**	**13180**	**14362**	**9.0**			

On average, across all 12 routes, fuel use in *Concept-B* was 9.0% higher than fuel use under *Concept-A*. The largest fuel use disparity was seen in Route 4, where *Concept-B* fuel use was 15.3% higher than *Concept-A* fuel use. The average company route in Route 4 used 15,235 pounds of fuel. Twelve of 16 subjects in *Concept-B* modified a company route, with the other four choosing a pre-exiting company route. The average fuel use of 14,445 pounds represents an improvement over the exiting company route fuel use average. However, the optimal route available in *Concept-A* involved a vertical solution where the most direct route was flown at 35,000 feet rather than 29,000 feet. At this higher altitude the fuel efficiency is better and fuel use drops 12,219 pounds. Fourteen of 16 subjects in *Concept-A* chose the fuel-optimal route, and thus the average fuel was a much lower 12,529 pounds.

The smallest differential was seen in *Route* 5, at 5.8%. *Route* 5 required a simple vertical solution, where a large lateral deviation was impractical because of the geographic extend of the area of severe turbulence. Thus most solutions (regardless of whether they successfully avoided severe weather as defined by the meteorologist) remained in narrow range, making the fuel numbers the most similar across subjects.

The standard deviation was larger (14.3% for *Routes* 1-6, and 10.1% for *Routes* 7-12) under *Concept-B* than *Concept-A*. Variation should be less under *Concept-A* since more subjects are choosing the same solution, as opposed to *Concept-B* where each subject is constructing a unique (albeit similar) solution.

Some results for *Concept-A* are mildly surprising. One would expect to see numbers like those of *Route* 8, where under *Concept-A* there was no variation in fuel numbers. All 16 subjects had identical fuel numbers, and presumably they would all chose the AWIN-generated optimal (for fuel and weather avoidance) route. Yet Route 8 was the only *Route* with zero variation. In fact, not all subjects in the *Concept-A* condition chose the AWIN-optimized route, thus accounting for the non-zero variation in fuel numbers.

The results of Analysis of Variance for *Routes* 1-6 and *Routes* 7-12 are given in Table 34. The first thing to note is that the main effects of *Weather Case* and *Route* are significant. This is not surprising since, as mentioned above, fuel use is dependent on the distance between city pairs, and that distance varies greatly between *Route*.

Table 34. Table of Variance for fuel use.

Source of Variability	DOF	F-ratio (P) Routes 1-6		F-ratio (P) Routes 7-12	
Concept	1	**F = 166.5**	**(p < 0.001)**	**F = 87.8**	**(p < 0.001)**
Weather Case	1	**F = 801.7**	**(p < 0.001)**	**F = 649.2**	**(p < 0.001)**
Route (Wx Case)	4	**F =154.9**	**(p < 0.001)**	**F = 416.6**	**(p < 0.001)**
Subjects (Concept))	30	F =1.05	(p < 0.403)	**F = 2.13**	**(p < 0.002)**
Concept x Wx Case	1	F = 0.710	(p < 0.401)	F = 0.72	(p < 0.399)
Subject(Concept) x Wx Case	30	F = 0.77	(p < 0.800)	F = 1.074	(p < 0.380)
Concept x Route (Wx Case)	4	**F = 5.18**	**(p < 0.001)**	F = 1.96	(p < 0.105)
Error	120				

The effect of *Concept* is strongly significant (F = 166.5, p < 0.001 for routes 1-6; F = 87.8, p < 0.001 for routes 7-12). Subjects in *Concept-A* where given an automatically generated route that optimized for fuel while avoiding hazardous weather. Of the 192 trials under *Concept-A*, subjects chose the fuel-optimal route 170 times (see Table 29). Subjects in *Concept-B* were required to construct their own route while trying to minimize fuel use and avoid hazardous weather. Since they also had to integrate the weather picture, subjects tended to use a wider margin when routing around weather, adding to distance flown and hence fuel use. *Concept x Route* is significant for *Routes* 1-6. Of particular interest is that the fuel use difference between *Concept-A* and *Concept-B* is smallest for Route 5, which is also the shortest route and has only one weather hazard. Figure 11 illustrates the average fuel use for each of 12 routes, broken out by *Concept*.

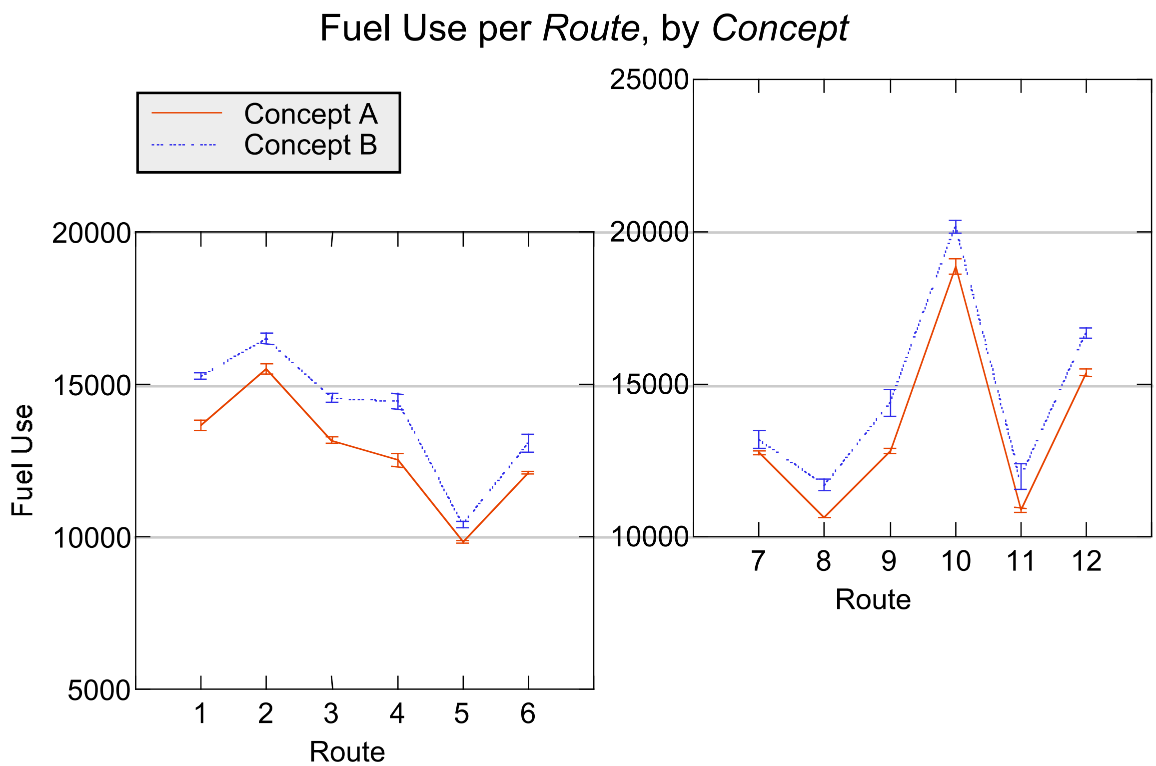


Figure 11. Fuel use on each route, under *Concept-A* and *Concept-B*.

5.5 *Planning Time*

Subjects conducted two blocks of six trials each. When subjects were conducting the *Concept-A* block, they were told they had one hour to complete the six trials. It was not a hard limit, but rather a way of setting expectations and discouraging subjects from spending too much time on off-task exploration of the tool. In the *Concept-B* block of six trials, subjects were told they had an hour-and-a-half. If subjects ran over the time, they were allowed to continue. The goal was to discourage them from excessive "tweaking" as they modified routes, which could add large amounts of time without improving the final route selection. Table 35 summarizes the descriptive statistics of planning time for *Route* 1-6 and *Route* 7-12.

Table 35. Summary of descriptive statistics of planning time, in seconds.

Metric	Concept A		Concept B.		Overall	
Routes	1-6	7-12	1-6	7-12	1-6	7-12
N of cases	90	95	96	96	186	191
Minimum (sec)	29	30	51	48	29	30
Maximum (sec)	1133	1094	2114	1077	2114	1094
Median (sec)	184	207	355	382	263	294
Mean (sec)	244	255	457	402	**354**	**329**
Std. Error	21	20	34	21	22	15
Std. Deviation	198	192	331	209	294	213

Subjects in *Concept-A* averaged 244 seconds for *Route* 1-6 and 255 seconds for *Routes* 7-12. Subjects in *Concept-B* averaged 457 seconds for *Route* 1-6 and 402 seconds for *Routes* 7-12. The minimum planning time spent in *Concept-A* was 29 seconds, and in *Concept-B* was 48 seconds, both of which are very low considering the tasks subjects were asked to perform. While the average times seem more than reasonable, it would be a concern if there were an excessive number of trials where subjects did not seem to make much of an effort (especially in *Concept-B*, where understanding the weather situation, reviewing the company routes for intersections, and modifying a route should take more than 45 seconds). Figure 12 illustrates the distribution of planning time per trial by grouping all planning times into bins within a minute. There were four trials in *Concept-B* that finished in one minute or less, or about 2% of the 192 *Concept-B* trials. Thus, it can be assumed that subjects made a serious effort in completing the trials.

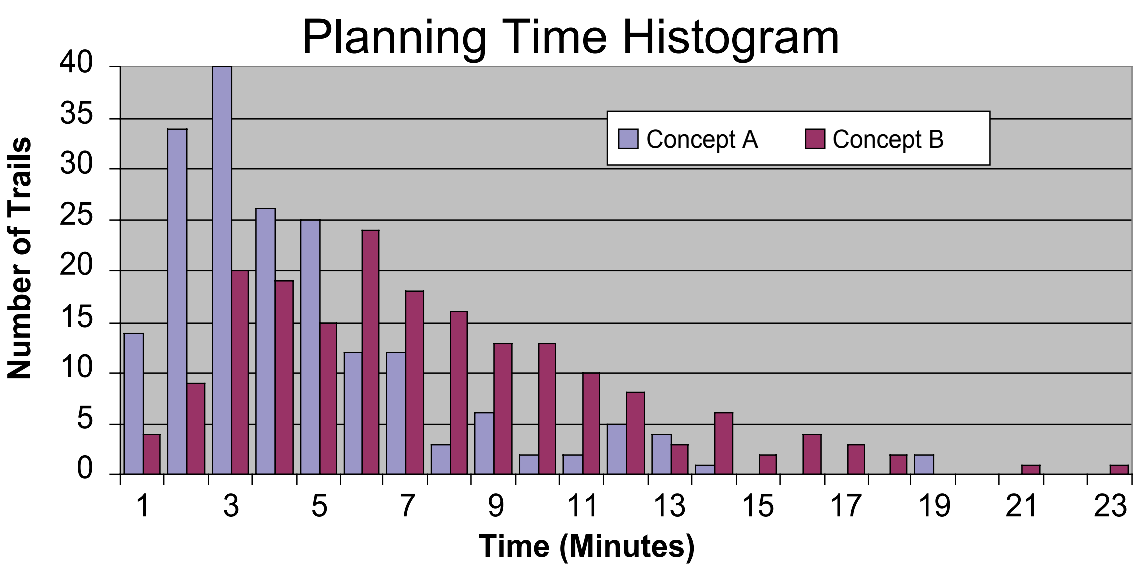


Figure 12. Frequency of planning times across all trials, in minutes.

Clearly the distribution of the *Concept-A* trials is grouped at shorter times than the *Concept-B* trials. Half (50%) of the 189 trails (three trials had no data) in *Concept-A* were completed in under 191 seconds, while it took 381 seconds for half of the 192 *Concept-B* trials to be completed. Most (90%) of the *Concept-A* trails were finished in under 522 seconds, while it took 758 seconds for most (90%) of the *Concept-B* trials to be finished. The longest *Concept-A* trials were just under 1133 seconds (just under 19 minutes), and the longest *Concept-B* trial (the only planning time not shown on Figure 12) was 2114 seconds (just over 35 minutes).

The results of Analysis of Variance for *Routes* 1-6 and *Routes* 7-12 are given in Table 36. The main effects of *Weather Case* and *Route* are not significant.

Table 36. Table of Variance for planning time.

Source of Variability	DOF	F-ratio (P) Routes 1-6		F-ratio (P) Routes 7-12	
Concept	1	**F = 43.5**	**(p < 0.001)**	**F = 42.3**	**(p < 0.001)**
Weather Case	1	F = 1.50	(p < 0.227)	F = 0.80	(p < 0.372)
Route (Wx Case)	4	F = 1.17	(p < 0.327)	F = 1.40	(p < 0..239)
Subjects (Concept))	30	**F = 3.22**	**(p < 0.001)**	**F = 3.15**	**(p < 0.001)**
Concept x Wx Case	1	F = 2.40	(p < 0..124)	F = 1.53	(p < 0.219)
Subject(Concept) x Wx Case	30	**F = 2.26**	**(p < 0.001)**	**F = 2.72**	**(p < 0.001)**
Concept x Route (Wx Case)	4	F = 0.33	(p < 0.858)	F = 0.95	(p < 0.441)
Error	120				

The effect of *Concept* is strongly significant (F = 43.5, p < 0.001 for routes 1-6; F = 42.3, p < 0.001 for routes 7-12). Planning time averaged 354 second for *Routes* 1-6 and 329 seconds for *Routes* 7-12. The average planning time for *Concept-A* (244 seconds for *Routes* 1-6, and 255 seconds for *Routes* 7-12) is significantly less than the average planning time under *Concept-B* (457 seconds for *Route* 1-6, and 402 seconds for *Routes* 7-12). The increased planning time in *Concept-B* can be attributed to the additional tasks subjects must perform in *Concept-B* that are automated in *Concept-A*: integration of route and weather information, and generation of a fuel-efficient route that avoids weather. Mentally integrating weather and route information across two screens slows down the process, as does manual modification of routes to avoid weather. Figure 13 illustrates the planning time per Route under both *Concept-A* and *Concept-B*.

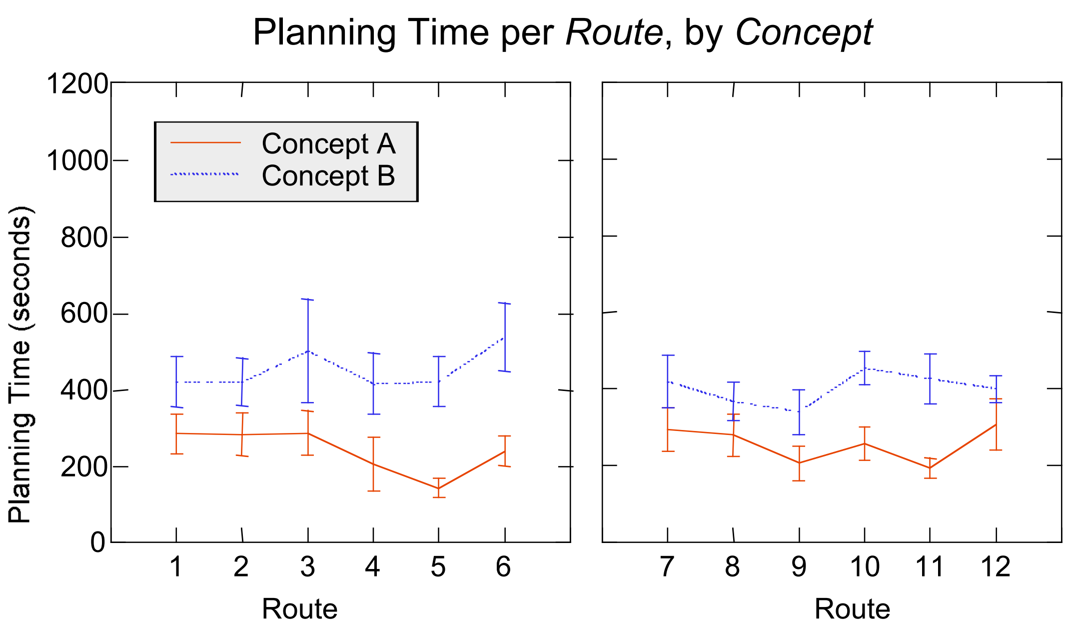


Figure 13. Planning time for each *Route*, under *Concept*-A and *Concept-.B*

5.6 *Weather Sources Accessed*

Subjects had access to MOCK during all 384 trials, under all conditions. Via the access logs, we were able to determine how many times a link was accessed through the MOCK page. Figure 14 illustrates how many times a particular MOCK page type was "hit" (accessed via a link). The totals shown in Figure 14 for each page is color-coded as the sum of the hits in each of the four weather cases. No *Weather Case* seemed to have a disproportionate share of the total hits for a MOCK page type. National radar summaries were by far the most popular MOCK page type, accounting for 525 of 2310 hits (22.7%). Infrared satellite was the next most popular at 256 hits (11.1%). No other page had more than 7.7% of the hits.

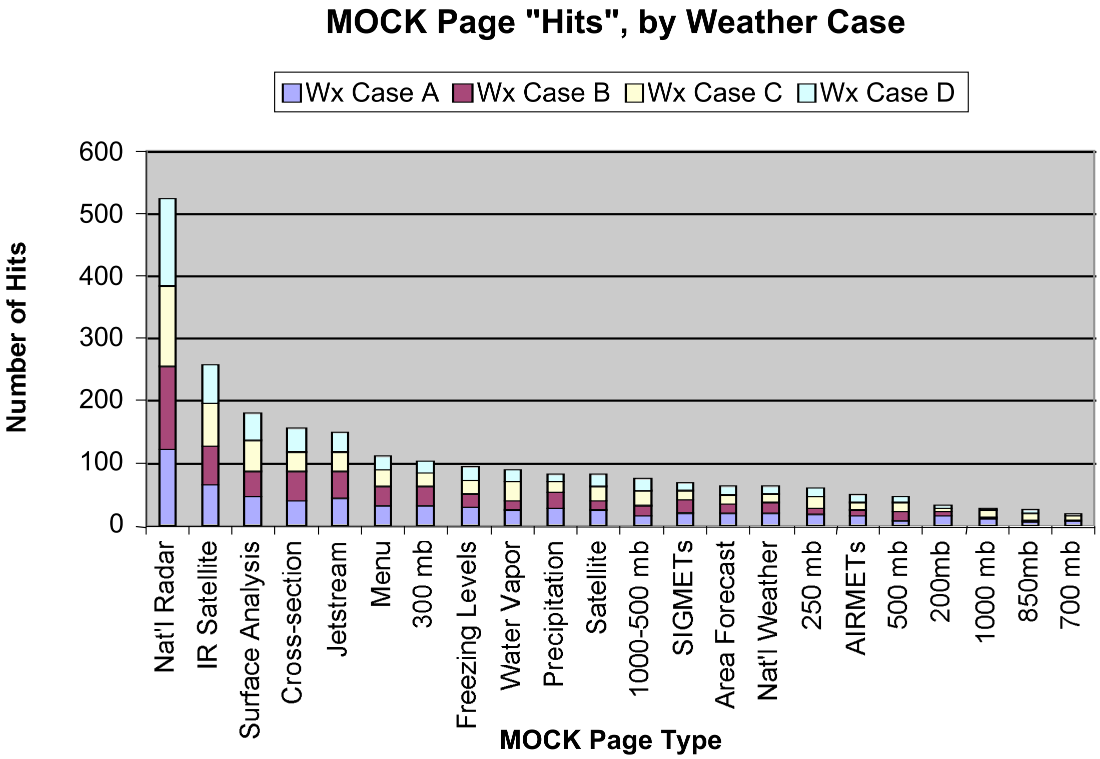


Figure 14. Number of hits per MOCK page, by *Weather Case*.

Concept-A had weather information integrated into the tool via weather polygons, while *Concept-B* did not. Figure 15 compares the number of hits per MOCK page type by *Concept*. Of the 2310 hits across all trials, 1415 (61.2%) of the hits occurred during trials under *Concept-B*. The relative proportions of hits per each MOCK page type remained the same; subjects simply accessed the pages more frequently during *Concept-B* trials, presumable since they had to mentally integrate the weather information in MOCK with the route information in the tool.

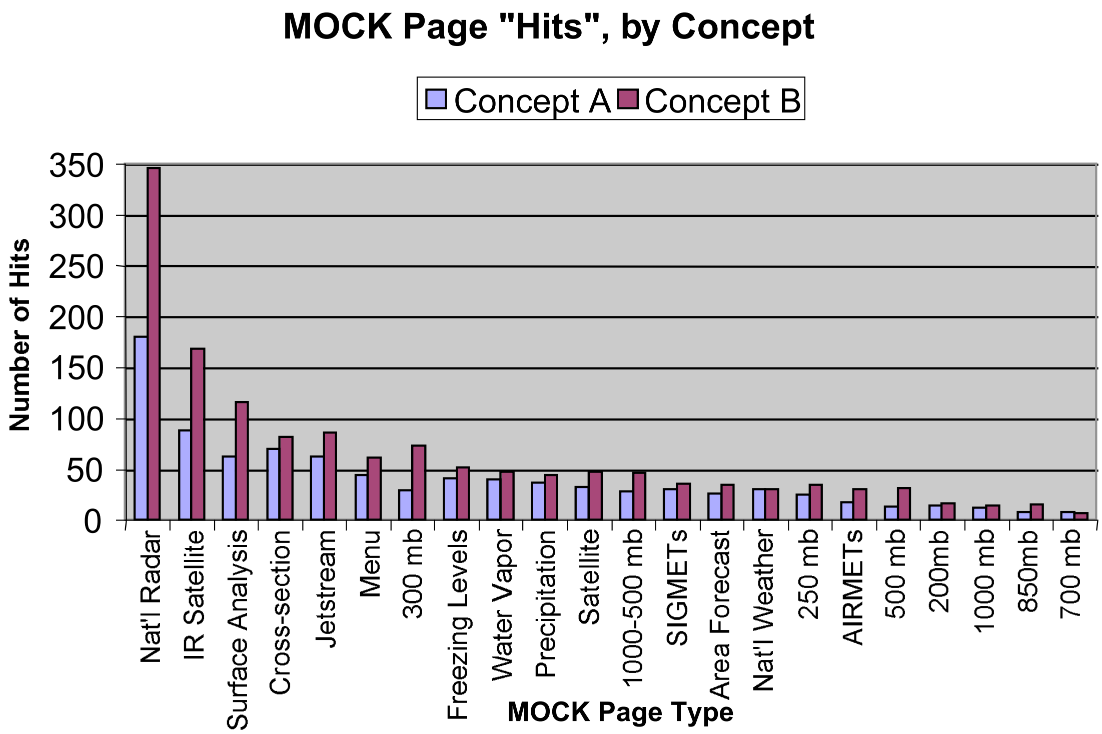


Figure 15. Number of hits per MOCK page, by *Concept*.

5.7 *Situation Awareness Probes*

Subjects were given five minutes before every set of three trials (city-pairs) to study the weather information found in MOCK. Since they did not know the city-pairs in the upcoming trials utilizing the weather case, the five-minute weather familiarization session was designed to give subjects time to create a general mental picture of the weather situation. Accordingly, when the weather familiarization session ended, subjects were asked a single question to probe their level of situation awareness. They were not allowed to look back at MOCK to find the answer. Each question was specific and had a definite correct answer. Table 37 lists the question for each weather case, as well as the number of subjects who answered correctly, incorrectly, or indicated that they didn't know (including those who left it blank).

Table 37. Summary of situation awareness probes statistics.

Weather Case	Question	Correct	Wrong	Do Not Know	% Correct over all 32 subjects
A	Is it raining in the Bismarck, North Dakota area?	18	12	2	56 %
B	Assuming a cruise altitude of 28,000 feet, would a route passing through the Minneapolis-St. Paul area in 3 hours encounter significant wind shear (speed or direction)?	10	11	11	**31 %**
C	Assuming a cruise altitude of 23,000 feet, would a route passing over the San Francisco area in 4 hours encounter a moderate layer of high relative humidity (>= 70%) with temperatures from 0 to –20C?	18	5	9	**56 %**
D	Are Dallas/Fort Worth area RADAR reflectivities currently in the 40-50 DBz range (or lower)?	18	8	6	**56 %**
	Overall Percentage:	**50 %**	**28 %**	**22 %**	

For *Weather Cases* A, C, and D, 18 of 32 subjects answered the question correctly (with the rest answering incorrectly or indicating that they did not know the answer). For *Weather Case* C, however, only 31% answered the question correctly. The questions were designed to require specific types of knowledge: (1) either of current conditions in question A and D, or forecasted information as in the case with question B and C, and (2) specific geographic locations. Many subjects commented that they did not remember the weather for the specific locale asked for in the questions. Given the difficulty of the questions, the overall 50% correct ratio versus 28% incorrect and 22 % "don't know" is a good indication that subjects, in general, were making good use of the weather familiarization session to gain an appreciation of the weather situation.

5.8 *Questionnaires*

5.8.1 Post-Block

Subjects conducted the 12 trials in two blocks of six trials each. One block was conducted under *Concept-A*, and the other block was conducted under *Concept-B*. Order of the blocks was randomized, as was the *Route* order within a block. After each block, subjects were asked to fill out a short "post-block" questionnaire.

The first question asked subjects if they felt that they had a good situation awareness (SA) of the weather situation. Table 38 details a comparison between subject assessment under *Concept-A* and *Concept-B*.

Table 38. Comparison of post-block subject situation awareness assessment.

Question 1. I feel that I had a good awareness of the weather situations.		
	Concept-A	**Concept-B**
N	32	31
Minimum	3	1
Maximum	7	7
Mean	**5.2**	**4.4**
1-factor ANOVA	F = 10.06 (p<0.003)	

Subject written comments are summarized, organized by *Concept*.

Concept A:

- Going back and forth from AWIN tool to MOCK made it difficult to translate weather information (S2, S17).
- Not enough SA time to get a complete picture of weather if you don't know what area you will need to look at from the trials (S1, S14, S26).
- MOCK was informative (S19), has all the information you need (S10), or even provides too much information (S25).
- Some weather information was missing or hard to use (specifically SIGMETs, AIRMETs and PIREPs) (S2, S19, S22).
- The software (S7) and especially the weather overlays (S29, S30, S31) were extremely useful.

Concept B:

- Some weather information was missing or hard to use (specifically SIGMETs, AIRMETs and PIREPs) (S1, S12, S14, S22, S26).
- Would like weather built into program (S1), more graphics (S9), or weather overlays (S31).
- Going back and forth from AWIN tool to MOCK made it difficult to translate weather information (S2).
- Too much weather information (S19, S37).
- Some subjects not familiar with weather products to interpret them effectively (S16).
- Felt working with MOCK was easier (S25), and that they read the weather data accurately (S30).

Subjects felt they had a better situation awareness under *Concept-A* (average ranking 5.4 on a 7 point scale) than in *Concept-B* (4.4), with the effect being significant (F=10.06, p<0.003). Subject comments identified some missing or difficult to use weather information in MOCK, most specifically text SIG/AIRMETs and PIREPs. Other subjects felt that there was enough information, but would have like to see more graphical weather depictions. Subjects would have liked to know what weather information to focus on during the weather familiarization session, but that would have run contrary to the needs of the experimental setup (see 4.1 on trial independence). The principal weather-related advantage of *Concept-A* was the polygon overlays.

The second question on the post-block questionnaires was designed to assess whether subjects had all the information they needed to complete the trials under the two conditions. Table 39 details a comparison between subject assessment under *Concept-A* and *Concept-B*.

Table 39. Comparison of post-block subject information completeness.

Question 2. I had all the information I needed to make a good routing decision		
	Concept-A	Concept-B
N	32	31
Minimum	3	2
Maximum	7	7
Mean	**5.6**	**4.4**
1-factor ANOVA	F = 17.57 (p < 0.001)	

Subject written comments are summarized, organized by *Concept*.

Concept-A:

- The information on the AWIN system was excellent (S2, S25), flight overlays (S13, S30, S31) and profile view (S7).
- MOCK was good (S11, S13, S25, S29), but some weather information was missing or hard to use (specifically SIGMETs, AIRMETs and PIREPs, constant pressure analysis, charts and other significant planning charts) (S2, S11, S14, S19).
- "Some weather seemed unavoidable, but the overlays helped out a lot" (S30).

Concept-B:

- MOCK was good (S2), but some weather information was missing or hard to use (specifically SIGMETs, AIRMETs and PIREPs, constant pressure analysis, charts and other significant planning charts, echo tops for convective weather) (S2, S26, S21, S32).
- *Concept-A* trials were easier with polygon overlays (S4, S30, S31). Weather information was there (in MOCK), but it was hard to "draw" on the route (S29).
- Need better weather graphics (S9, S24, S25), AIRMETs and SIGMETs in map form (S21).
- Time stamps were sometime not in synch (S3, S7).
- Need better information for a pilot. Info given better suited to meteorologist (S26).

Subjects felt that they had enough information in *Concept-A* (average rating 5.6 on a 7-point scale), more than they had in *Concept-B* (4.4). The effect of *Concept* was significant in a one-factor ANOVA analysis (F=17.57, p < 0.001). Subjects identified the overlays and the profile view as the most helpful. Subjects in *Concept-B* felt it was hard to integrate disparate weather and route information. Interestingly, one subject felt that some of the weather "seemed unavoidable." All routes in the experiment had weather that could be avoided, either via a lateral or vertical modification of a company route.

There was one question that appeared only on the *Concept-A* post-block questionnaire as it pertained to the weather polygons found in the tool under *Concept-A* only. Subjects were asked if they felt confident in the hazards boundaries and severities as defined by the company meteorologist. The intent was to assess the degree of fit between the subject's view of the weather, and the meteorologist's view. Table 40 details subject response.

Table 40. Subject assessment of polygon definitions in *Concept-A*.

Question 2. I feel confident that the hazards accurately captured the boundary and severity of the weather hazards		
	Concept-A	**Concept-B**
N	32	NA
Minimum	3	NA
Maximum	7	NA
Mean	**5.4**	NA

Subject written comments are summarized:

- Subjects liked the fact that the weather moves and you can see where it's going. (S1); the animation was excellent in diagnosing weather phenomena coverage (S19).
- Cross-checking with the MOCK screen the information seemed inaccurate (S2), or hard to tell if they were accurate (S12).
- "Switching between screens seemed to increase the chances of misreading or misplacing weather information onto the AWIN screen" (S2).
- "Some were quite broad (S5, S11, S19, S30, S31), yet others had precision uncommon in meteorology" (S5). The boundary could be safer (S13). Overlays were nice (S29).
- Because much of the information is forecasted and a meteorologist had to do all the painstaking input of data, there can and will be errors involved (S7, S22). However, this method is far better than what is currently in use (S7).
- The vertical view helps you plan much quicker (S10).
- Mixed with radar and other information, easy to plan (S11).

Subjects rated the goodness of polygon definition as 5.4 on a 7-point scale, indicating that most subjects felt the polygon definitions were true to the weather phenomena. Some subjects indicated that they had some trouble verifying weather definitions via cross-checking with MOCK, and some subjects felt the boundaries were too large. Subjects indicated that they felt the animation feature was useful in diagnosing weather coverage. Finally, subjects indicated their support of using polygons, citing the time savings as well as an improvement over current practice.

The next question (number four on *Concept-A* post-block, and question 3 on the *Concept-B* post-block) asked subjects if they were confident in the final route they had selected (*Concept-A*) or constructed (*Concept-B*). Table 41 details a comparison between subject assessment under *Concept-A* and *Concept-B*.

Table 41. Comparison of post-block subject confidence in route selection.

Question 4(A)/3(B). I am confident that I selected/constructed fuel-efficient routes that avoided hazardous weather.		
	Concept-A	Concept-B
N	32	31
Minimum	4	3
Maximum	7	7
Mean	**5.9**	**5.3**
1-factor ANOVA	$F = 5.09$ ($p<0.031$)	

Subject written comments are summarized, organized by *Concept*.

Concept-A:

- As long as the dispatcher accurately checks significant weather systems the performance charts, and other significant data, the system seemed to alleviate most of the work involved (S2).
- I chose the route that avoided all weather (S7).
- I selected the route with the least amount of fuel (S13).
- I chose the route recommended by the computer (S22) [Note: there was no “recommended route”].
- The route comparison tool was invaluable (S19). The overlays with winds and hazards were useful (S29, S30).
- Having pre-determined altitude made it easier (S30).

Concept-B

- Mostly fuel efficient by trial and error as well as more direct routes (S12).
- Did not go through hazardous weather (S13), I feel I did an adequate job (S29), for the most part (S31).
- Being able to understand exactly what the forecast would have helped as well as winds for specific altitudes” (S14) [note: winds were available]. I felt I was guessing at times, not sure of weather tops and turbulence (S1).
- One flight I would have cancelled (S16). Some weather seemed to be unavoidable without canceling the flight (S30).
- I cut the thunderstorm close a couple of times. Crew may have had to divert more than once (S19).
- I found frustration with my lack of ability to understand data to make good choices especially related to time (S22). Difficult to analyze AIRMETs and

SIGMETs from other parts of the country as my knowledge of airport identities is limited to major airports (S2).

- Not sure weather in many circumstances. If planning is done in graphical format, all weather should be as well (S26).

Subjects felt very confident about the route they selected (5.9 in *Concept-A*) or constructed (5.3 in *Concept-B*). The effect of *Concept* was significant in a one-factor ANOVA analysis (F=5.09, p<0.031). Subject comments reflected some of the various strategies they used in selecting route in *Concept-A*, ranging from the route that avoided all weather to the route with least fuel. One subject wrote that they selected the recommended route, but there is no official "recommended route" – presumably the subject meant the AWIN-generated optimized route. As one subject pointed out, one must be careful to check the significant weather before feeling confident about selecting a route. Subjects in *Concept-B* had a harder task, and the lower score and some of the comments support their frustration in having to mentally integrate weather and route information. They were less sure of the weather. Again, some subjects felt at least one flight should have been cancelled rather than trying to find a route through the weather. The meteorologist had determined that all *Route* were flyable, so this represents a difference of opinion and/or strategy by the (two) subjects. It is possible that they did not consider possible alternatives (*e.g.,* a vertical solution, as some subjects did not seem to realize this was an option).

The final question asked subjects if they were confident in the final route they had selected was safe. Table 42 details a comparison between subject assessment under *Concept-A* and *Concept-B*.

Table 42. Comparison of post-block subject confidence in route safety.

Question 5(A)/4(B). I feel that the routes I selected were safe.		
	Concept-A	**Concept-B**
N	32	31
Minimum	3	4
Maximum	7	7
Mean	**6.0**	**5.4**
1-factor ANOVA	F = 7.05 (p<0.013)	

Subject written comments are summarized, organized by *Concept*.

Concept-A

- I felt the routes were very safe (S29). They are safe as long as the dispatcher accurately cross checks and verified the weather information (S2).
- I always went with the route of least threat (S5). I avoided all severe weather, at least as predicted and forecasted (S7).
- I am not sure it was safe (S12). Some routes were cutting it close, but that is for ATC to decide (S14).
- The very last route I chose may not have been the safest, but I feel it was not dangerous in any way (S30).

- Keeping flight safe was easy, except when boundary conflicts among polygons arose (S19).
- I would like to see charts/info used to generate the hazard areas (S22). Information provided by AWIN was good but it should be easier to look it up (back it up) in MOCK (S25).

Concept-B

- Not sure of weather (S26), it might change en-route (S8).
- The selected routes were safe and efficient (S29). Most of the selections were comfortable (S8). I did not go through hazardous weather (S13).
- I am not sure they are safe (S12). I was guessing at times (S1). I was frustrated at my lack of ability to understand data and make good choices (S22).
- Some weather was unavoidable (S30).

Subjects on average felt very confident that they routes chose were safe for *both Concept-A* (6.0 on a 7-point scale) and *Concept-B* (5.4). The effect of *Concept* was significant ($F=7.05$, $p<0.013$) in a one-factor ANOVA analysis. Subject comments in *Concept-A* reflect their feelings of confidence. Subject comments in *Concept-B* reflect the lower confidence, where some subjects reported that they felt that they were guessing at times. But in both conditions, overall subject confidence in the safety of the routes was high.

5.8.2 Post-Experiment Comparative

The first section of the Post-experiment questionnaire asked subjects to rate, on a single 7-point scale, *Concept-A* and *Concept-B* on several dimensions. Table 43 lists the subject ratings for each dimension. The ratings are on a seven-point scale. A one-factor ANOVA test was done on each rating to see if the effect of *Concept* was significant.

Table 43. Comparison of *Concept-A* and *Concept-B*, ratings on a 7-point scale, with significance tests.

Question	Concept A				Concept B				P<
	N	Min	Max	Mean	N	Min	Max	Mean	
1. How well does the tool enhance safety?	29	2	7	**6.0**	27	2	7	**4.1**	.01
2. How well does the tool assist you in comparing flight plans to assess fuel efficiency?	28	2	7	**6.1**	28	1	7	**5.8**	.58
3. How well does the tool assist you in detecting when a flight plan intersects hazardous weather?	29	3	7	**6.3**	27	1	7	**3.6**	.01
4. How difficult was it to read and understand the information on the tool's display?	29	3	7	**6.1**	27	3	7	**5.7**	.08
5. How well where you able to assess the severity of weather?	29	3	7	**6.0**	27	1	7	**3.8**	.01
6. How comfortable are you in ability to use this tool for flight routing?	29	4	7	**6.1**	26	2	7	**5.0**	.01
7. How valid were the scenarios?	29	2	7	**5.8**	27	1	7	**5.8**	.01

Question 1 asked subjects how well they felt that the tool enhances safety. A summary of subject comments and observations is as follows:

- Concept B gives more freedom to plan routes around weather as you can see fit. A constrains your decisions to a few select choices (S11). It would be great to be able to have both concepts going at the same time (S1, S11, S12).
- Concept A was a tremendous enhancement at helping to avoid bad weather situations provided it accurately described the weather systems (S2, S3, S8, S22, S30). You need weather graphics with route modifications (S9, S10, S24, S31).
- Concept A leaves little doubt for the dispatcher as to what a meteorologist is expecting, much more user friendly and dynamic for each individual flight plan (S5, S7, S13, S15).
- Concept A is a great tool to estimate what altitude is necessary and if the flight can even be completed safely (S7, S17, S29).
- Concept A had problems with boundary lines. A Dispatcher should not have to trust polygons as much as A forced us to do (S19). Using Concept A you can get some conflicting information (S25).
- Concept B requires too much thinking and analysis to be useful to most pilots (S26).

Subjects strongly preferred *Concept-A* (average rating 6.0 on a 7-point scale, where *Concept-B* was rated as 4.1), feeling that the weather overlays reduced workload and made it easier to avoid hazardous weather. The effect of Concept was significant ($F=16.50$, $p < 0.001$). Some subjects felt that the ability to modify routes was good and should be incorporated into a final tool that also included overlays. Some subjects expressed concern with the way boundary lines were drawn, while others felt that *Concept-B* required to much analysis to be useful.

Question 2 asked subjects how well the tool assists them in comparing flight plans to assess fuel efficiency. A summary of subject comments and observations is as follows:

- The route comparison tool was very useful in assessing fuel (S5, S8, S17, S18, S24, S31) in both conditions (S3, S10, S11, S12, S22). Does it all for you (S7, S16, S26).
- Nicest feature of the program, along with weather polygons (S2).
- Creating waypoint was paramount in determining course (S19).
- B was easier to assess fuel efficiency because the flights (especially their altitudes) could be altered to increase fuel efficiency (S13, S30).

Subjects felt there was little difference between *Concept-A* (average rating 6.1) and *Concept-B* (5.8) since both tools calculated fuel consumption via the Route comparison tool. The effect of *Concept* was not significant ($F = 0.31$, $p < 0.580$). Two subjects felt that ability to change altitudes in *Concept-B* allowed for greater efficiency.

Question 3 asked subjects how well the tool assists them in detecting when a flight plan intersects hazardous weather. A summary of subject comments and observations is as follows:

- Both the vertical and horizontal depictions greatly enhanced my weather situational awareness. My only concern is that dispatchers might get lazy about double checking the polygons to the actual MOCK weather (S2).
- The depiction of weather in *Concept-A* was helpful (S3). It is a great tool for viewing flight hazards (S3, S5, S7, S8, S12, S16, S20, S21) and hazards intersecting weather (S22, S24), as long as the person who inputs the weather has some degree of accuracy (S5, S14, S25, S30) and familiarity of the software (S5).
- There was no weather information or help in *Concept-B* (S9, S11, S12, S26, S29, S30). Without the hazards shown one might overlook an area of bad weather because there is so much to look at (S10). You have to be really good in the weather otherwise it will take you a long time to gather all the data, analyze it and then apply (S13). Concept A is definitely better for those of us who are more visual (S10).
- The animation feature with a 3D view is extremely helpful (S14), crucial (S19).
- Tool B functioned better for me because it allowed changes to flight plans and because I relied on MOCK which contains the actual reports and forecasts (S25).

The weather information contained in *Concept-A* allowed subjects to see routes intersecting weather, something not available in *Concept-B*. Subject comments reported a wide acceptance for weather polygons, but several subjects warned of complacency in trusting the weather polygons without checking their accuracy. One subject preferred *Concept-B* precisely because it forced him or her to work with the raw weather data. *Concept-A* reduced workload, and subjects commented specifically on the usefulness of the vertical and lateral views, as well as the animation feature that allowed subjects to get an appreciation of the movement of forecasted weather. The effect of *Concept* was significant ($F = 28.56$, $p < 0.001$), as there was a large difference in rating between *Concept-A* (6.3) and *Concept-B* (3.6).

Question 4 asked subjects how difficult it was to read and understand the information on the tool's display. A summary of subject comments and observations is as follows:

- Very User friendly (S2, S5, S25, S30), easy to use (S7, S10, S12, S19, S22), easy to understand (S24). Very difficult to "mess up" (S5).
- Some of the overlapping routes were hard to pick (S3).
- *Concept-B* was difficult because you constantly had to refer back to charts to get an idea of what was going on with the weather throughout your route of flight (S7). Concept A had it all laid out for you. (S7, S13).
- *Concept-B* was less cluttered (S11). Concept A was difficult because of having to impose the weather onto the route display (S14).
- The user should be allowed to customize window sizes (S19).
- Animation was not expected at first, I had to learn it myself (S26).
- Not bad (S18), not difficult in either concept (S29).

Subject comments overwhelmingly felt that the system was user friendly, easy to understand, and easy to use. The effect of *Concept* was marginally not significant ($F = 3.24$, $p < 0.084$). One subject felt *Concept-B* was less cluttered and another felt *Concept-A* was difficult to use because of the weather being superimposed on the route display. Two subjects commented that they liked having all the information in *Concept-A* laid out for them.

Question 5 asked subjects how well they were able to assess the severity of weather. A summary of subject comments and observations is as follows:

- Using two computers at the same time in *Concept-B* is not efficient (S1, S12, S25), and makes it hard to remember exact locations (S2).
- The detail boxes helped enormously in *Concept-A* (S2).
- Both concepts would have been nicer if you could overlay raw weather information on the display (S2, S12).
- It was much easier in *Concept-A* (S7, S8, S25, S29), all you had to do was look at the graphics (S9, S10, S12, S14, S22), assuming that the weather graphics are accurate (S25, S29).
- Concept B you had to use more guessing (S9, S10, S24), or figuring it out (S14, S23, S26). But doing the research gives you a lot of extra information (S13).
- the basics were there, although access to echo tops of convective cells would be highly useful (S5) Reading text weather was confusing (S10). it was sort of tough to assess severity of the precipitation cells. Radar summary charts could be helpful for that (S24).
- Assessing the severity of weather was fairly easy for both systems (S11, S30).

Subjects felt that assessing severity was easier in *Concept-A* than in *Concept-B*. The effect of Concept was significant ($F = 38.65$, $p < 0.001$) as there was a large difference in rating between *Concept-A* (6.0) and *Concept-B* (3.8). In *Concept-B* several subjects reported that they felt like they were guessing, with some subjects requesting more information in MOCK. Two subjects felt that their ability to assess the severity were about equal in both conditions.

Question 6 asked subjects how comfortable they were in their ability to use this tool for flight routing. A summary of subject comments and observations is as follows:

- I prefer to have weather done for me, rather than modifying routes on weather data (S3). I like Concept- A better because of the ease in avoiding hazardous weather (S22).
- it is simply a matter of understanding the hazards which impact operation (S5).
- The software does not need much explaining to get started (easy to use (S12, S18)), thus making it a very good investment (S2, S7).
- I think that using this tool for flight routing will put very good position because it help a lot about routing information and weather (S8).
- I feel safer using *Concept-A* than *Concept-B* (S10, S29, S30). I don't like the lack of info presented in *Concept-B* (S26).

- I would rather just make my own flight plan (S13), I like having more freedom (S11). I am comfortable with *Concept-B* because it reduces the risk of human error in displaying hazardous weather at due times along your route in comparison with the aircraft position (S14).

Most subjects felt comfortable using the tool to do flight planning. The effect of *Concept* was significant ($F = 12.44$, $p < 0.002$). Many subjects expressed preferences *for Concept-A* (6.1 rating) because it presented processed weather hazards, making it easy to assess the impact of weather on their flights. Three subjects expressed a preference for *Concept-B* (rating 5.0) since they felt it afforded more freedom (modifying routes) and forced them to consider the raw weather data.

Question 7 asked subjects how valid they thought the scenarios were. A summary of subject comments and observations is as follows:

- They are both valid. This is because flights are affected each and every day by weather concerns. Also, we all know the fuel is one of the major costs for the airlines, so it's very to be fuel efficient- but safe at the same time (S10). Weather and fuel are only one of the many other factors that are considered while planning flights (S31).
- In my limited experience they seemed valid (S2, S12), or difficult to assess (S7).
- They were valid (S5, S24, S27) and very good (S19, S22), and realistic (S25). They seemed very valid due to the routes assigned along the weather at the assigned flight time (S24).
- It was real weather data (S13, S16, S18, S27). Need more weather data (S26).
- The scenarios were consistent with real operations (S29, S30), although ATC would have figured in (S29).

Subjects rated the validity of the scenarios equally in the two conditions, and the effect of *Concept* was not significant ($F = 0.033$, $p < 0.857$). The scenario's validity rested primarily on the fact that real data was used, and the routes were consistent with real operations. The weather came from actual recorded weather, and the company routes were selected from a database of actual routes flown by airlines.

5.8.3 Post-Experiment Trust

The second section of the post-experiment questionnaire asked subjects to rate their level of trust in various aspects of the information, the tool, and their confidence in their ability to perform the tasks asked of them. Table 44 lists the ratings subjects gave for each question, on a 10-point scale.

Table 44. Subject assessment of their level of trust and confidence.

Question	N	Min	Max	Mean
8. Overall, how do you rate your trust that the CONCEPT A flight planner will produce the most fuel-efficient route?	29	5	10	**8.1**
9. How do you rate your trust that the hazard polygons accurately represent the presence of weather hazards (*i.e.,* any significant weather that exists will be represented by hazard polygons within CONCEPT A)?	29	3	10	**7.6**
10. How do you rate your trust that the hazard polygons accurately represent the spatial extent of the weather (*i.e.,* 3D polygon boundaries encompass significant weather)?	29	3	10	**7.3**
11. How do you rate your trust that the hazard polygon severity accurately represents the severity of the weather?	29	3	10	**7.5**
12. Using the CONCEPT A Tool, how high is your self-confidence in selecting a fuel-efficient route that avoids hazardous weather?	29	3	10	**8.0**
13. Using the *Concept-B* Tool, how high is your self-confidence in constructing (via the manual manipulation of waypoints on a company route) a fuel-efficient route that avoids hazardous weather?	29	2	10	**6.8**
14. Rate the trust you had in the weather information used in the trials?	29	3	10	**7.8**

Question 8 asked subjects to rate their level of trust in the *Concept-A* tool to produce the most fuel-efficient flight plans. They rated it as 8.1 out of a 10, displaying a high level of trust.

Questions 9, 10, and 11 related to polygon definitions. Subject trust in the polygon hazard definitions also rated highly. Subjects rated their trust (7.6) that all relevant hazardous weather, if it existed, would be displayed as polygons in *Concept-A*. They rated their trust in the boundaries of those weather polygons as 7.3, slightly lower but still high. Finally, the polygon severity rating earned a 7.5 trust rating. Thus subjects trusted the existence, extent, and severity of the polygons they were presented with in *Concept-A*.

Question 12 and 13 asked subjects to rate their level of confidence in their performance when selecting a fuel-efficient route that avoided hazardous weather, in *Concept-A* and *Concept-B* respectively. Self-confidence in their performance in *Concept-A* (8.0) rated much more highly than self-confidence in their performance in *Concept-B* (6.8). Table 45 summarizes a one-factor ANOVA analysis between the two ratings, showing that the effect of *Concept* was significant ($F = 6.12$, $P < 0.020$).

Table 45. Table of Variance for self-confidence.

Source of Variability	DOF	F-ratio (P)
Concept	1	**F = 6.12 (p < 0.020)**
Error	28	

Question 14 asked subjects to rate their level of trust in the weather information found in the trials. Subjects rated it highly as 7.8 out of 10.

Question 15 simply asked subjects if they had any comments on the general issue of trust. A summary of comments and observations follows:

- I would not trust the weather polygons without being able to verify them (S2, S23, S25, S29) with actual aviation products (S1).
- Overlaying raw weather information on the planning map would increase my trust greatly (S2).
- It requires trusting the meteorologist accuracy (S4, S24, S25). I am not a meteorologist so my answers may be less reliable than that of a meteorologist. (S7).
- How can I trust that my route will be legal? Some weather was so borderline, that I didn't trust it to have completely dissipated as expected (S5).
- Some of the MOCK data had incorrect timestamps (S7). I wasn’t sure about the airports named in weather reports (S9). I had no trust because of incomplete or poorly presented data. (S26).
- Weather is unpredictable (S10, S16), but I don't believe that a system could be any better for visual flight planning than concept A (S10). As long as the weather is updated the tool will be very trust worthy (S27).
- Its not a question of trust, but of available information in a easy-to-understand format (S12, S26).
- I question the fuel burn calculation, the ground speed factor. I question climb and descend time and fuel requirements (S31).

Subjects reiterated that their level of trust in the weather polygons is dependent on their ability to verify the information content. Similarly, subjects said that it would take time to assess if they could trust the polygons, and the meteorologists who creates them. Some inconsistencies within the data found in MOCK, while not destroying their trust (which still rated highly) in the information, had caused some doubts and as a result subjects increased their vigilance by trying to find multiple sources of the same information.

5.8.4 Post-Experiment Written Feedback

The final section of the Post-Experiment Questionnaire was an opportunity to ask subjects questions where they could write comments and give the experimenters free-form feedback. We used the opportunity to ask questions on the best and worst features of both current practice and the AWIN prototype. We asked questions relating to the validity, efficacy, and workload of the experiment itself. Finally, we asked questions designed to elicit feedback for the next iteration of development.

Question 16 asked subjects to list the best and worst features of any current flight planning systems with which they were familiar. Table 46 lists a summary of user comments.

Table 46. List of subject assessment of the best/worst features of current flight planning systems.

Question 16. Please list best and worst features of any **current flight planning systems** with which you are familiar.	
Best	**Worst**
Weather displayed with the routes (S1, S4, S8). ability to display airspace limitations (S5)	Cost of updates and graphics (S1)
The use of a "rubber band" route option gives you a weather brief that applies to the route. (S1)	Lack of accuracy and fuel efficiency (S2). Doesn't optimize like the computer here did (S12)
All weather and planning is done manually myself, which assures it is done the way I want it. (S2)	Can only display things you ask, if you don't ask but its pertinent you still won't get it. Like a failsafe function. (S3)
Availability-extremely correct and constantly updated. Flight computer with flight data. (S3)	Some systems lack access to all available weather tools (S5).
Radar, doppler & intensities. (S6)	Winds aloft forecasts. (S6)
Allows me to get a big picture and manipulate my route (S12, S26) so I know exactly where I am going.(S12)	does not integrate weather and draw out your route on a map like *Concept-A* did.(S7)
DUATS (S24) - it gives a lot of information.(s13, S15). Factoring in winds aloft to calculate GS and ETA at fixes and destinations. (S7)	Many systems can't show you forecasted movement like the hazards can (S10). There are no 3D pictures (S23)
PIREPs, airmets, sigs, anything avail to the pilots (S18). the ability to get current metars, tafs, pireps and forecasts along my route. (S20)	The worst thing would be user friendly, I have had some problems with working with the software (S15). We should be able to customize (change) more settings. (S13)
pilot brief gives accurate information (S21)	coded information is sometimes difficult to understand (S20)
allows you to narrow down to a certain region when getting reports or forecasts (S25)	having to scroll through pages of forecasts/reports to find what you are looking for (S25, S27)
the best feature I have ever worked with was a system that gave a detailed, plain English description of anticipated weather along the route during the specified times (S30)	the worst system I ever encountered allow only origin to destination straight line flight, *i.e.,* no way points between MIA and PIT (S30). rigid VOR to VOR route process for IFR flight (S26)

Question 17 asked subjects to comment on the compatibility of the Honeywell prototype system with current systems. Is there anything they would need to "unlearn?" A summary of comments and observations is as follows:

- Not so much "unlearn" (S3, S19, S22, S23, S25, S26, S28) as just learn the different features of the program.(S1, S8, S10, S11, S13). Easy to use (S10, S19, S20, S27).
- I have not tried any other systems (S2, S4, S30, S31), but I think all of the tools were appropriate and useful (S9), there is nothing I would want to "unlearn" (S4, S30).
- With the addition of greater weather, ATC and airspace capability, this program would excel past those I have used (S5, S24). Would have to get better weather forecasting and updates to be truly usable.(S29).

- Hope to see more of this type of tech developed (S15). I feel that the Honeywell system could be easily integrated into a system all pilots. Dispatchers could use (S24).
- No. I think most of the things I did today were very valuable (S8, S18). Good system in terms of commercial travel (S6).
- It isn't in the greatest interest for light aircraft (S6). Some adaptation to the GA market could make it an invaluable GA tool (S19).

Subjects did not necessarily feel that there was anything to “unlearn”, rather that they would have to learn the new features of the Honeywell prototype planner. They felt the planner was easy to use, but would have liked to see more raw weather information integrated into the system.

Question 18 asked subjects how the reliability of weather information affects their willingness to plan a route through a hazard. A summary of comments and observations is as follows:

- Use with caution because weather is always changing (S1, S8, S16). if I cannot trust the weather info greater avoidance parameter are applied (S5, S23, S26, S30).
- The more reliable weather source, the more willingness I would have to plan through weather (S15, S20, S21, S24), very important (S18, S22, S28, S31), paramount to aeronautical decision making (S19).
- Having accurate information gives the “planner" confidence (S3, S4, S9, S11, S25) to route the aircraft through hazardous areas (although clearing them) and at the same time, be safe and fuel efficient (S4).
- I would not plan a flight with unreliable weather information. (S3, S13, S27). If there is a chance of a hazard, I will not attempt to get near it (S7, S12, S29). It’s the unforecast weather that brings up problems (S15).
- If it is unreliable I will refer to other sources to get an estimation (S7, S14). Only a few sources are reliable (S17).
- Unless you can visualize somehow where these hazardous areas are, your planning won't be very reliable (S10).

Subjects strongly cautioned against using unreliable weather information when flight planning, feeling that reliable weather information was crucial to good decision-making. When information is deemed unreliable, then subjects would either no plan the flight at all, or increase the safety margins around areas of potentially hazardous weather.

Question 19 asked subjects if there were rules of thumb they used in terms of when to consider a weather prediction certain and plan accordingly, and when to consider it uncertain and essentially ignore it? Related to Question 18, the intention of this question was to discern any heuristics subjects might employ when dealing with potentially unreliable information. A summary of comments and observations is as follows:

- None (S1, S4, S8, S10, S17, S18, S22, S23). I tend to rely upon the charts, but most of all a weather briefer with the FSS to advise me of potential weather development or dissipating weather (S15, S24, S30).
- I look at the overall weather picture (S2, S12), Consistency in reports and correlations with people/pilots currently flying and reporting (PIREPS) (S2, S3, S7, S12, S21, S25).
- If all weather data are in-line with each other, then I'd consider it accurate. (S2, S19, S20, S21). If it does not agree with a number of other sources then I usually will not consider it (S14).
- Past trends (S5, S6, S9), uncertainty among forecasters (S5, S11).
- Direction of movement and inspection of turbulence warnings and freezing levels (S6).
- A huge difference between high and low altitude charts is important (regarding severe weather) because upper air analysis determine where the surface systems are (S7).
- Once two or three sources that have been reliable in the past confirms the same thing I go with that (S16).
- Usually a TAF is pretty good. An Area Forecast is not worth the paper it is printed on, especially for icing. It can be useful for turbulence and convective weather, but a pilot learns to distrust it because of the ridiculously bad picture it often paints (S26).
- Any SIGMET or freezing level I avoid 100%, any other bad weather stay away (S27, S31).

Many subjects reported that they did not employ any rules of thumb, trusting instead weather information sources that they deemed reliable in the first place (and thus assuming that any information found within those sources must be reliable). Many subjects reported that they looked at the overall weather picture, and verified the reliability of weather information by looking for confirming or disconfirming information from alternate sources.

Question 20 asked subjects to list the best and worst features for flight planning with the Honeywell prototype flight planning system. Table 47 lists a summary of subject comments and observations.

Table 47. Subject assessment of the best/worst features of the Honeywell flight planning systems.

Question 20. Please list best and worst features for flight planning with the **Honeywell prototype flight planning system.**	
Best	**Worst**
Concept A: fuel efficient planning (S2, S9, S12, S13, S24) that avoids all hazardous weather (S1, S15)	Concept B: Using 2 computers to gather info is not that easy.(S1)
Concept A: The visual polygon weather depiction was very helpful (S3, S9, S10, S12, S13, S14, S21, S24, S26, S29, S30) for planning (S1) and situation awareness (S2, S4, S20)	*Concept-A*: polygons might lead to a lack of cross-checking with actual data (S2, S13, S22, S26, S29). If the polygons are constantly accurate, the dispatcher could begin to assure that it will happen at all times
Concept B: modifying routes (S4, S15, S16), enjoy the freedom to modify routes (S11)	Concept B: Not being able to differentiate between slightly modified routes on the main display screen.(S3)
Concept-A: directly linking company meteorologist with the dispatcher (S5)	Concept B: no ease of weather forecast in the same software (S4), does not include weather hazards (S10)
great ease of flight path selection (S6, S7, S21)	No airspace consideration (S5)
Concept-A: integration of weather and route information (S7, S8)	Concept A made me feel constrained and a bit anxious. No freedom to modify routes (S11, S31)
Route Comparison Tool (S14, S22)	Need more real world weather input, maybe overlay sigmets, airmets, freezing layers, etc., on map (S2, S12, S16, S17, S24)
Animation (S14, S18)	Could be more user friendly (S15, S17), but with the proper training I feel it would be a very effective tool (S15)
it provided good help for decision-making (S17). User interface (S19, S25, S27), features, and formats (S23)	*Concept-A*: the shapes of the weather should maybe be more radar depiction types (S18)
Vertical View and icons (S28)	Failure to adapt to other airframes (S19)

Subjects listed the weather polygons and integrated flight planning features of *Concept-A* as one of its best features. Conversely, the need to work between two screens *in Concept-B* was considered difficult, probe to error, and time-consuming. The *Concept-B* feature that allowed for *Route* modification was considered a good feature, as was the Route Comparison Tool, animation of polygon hazards, and the vertical view. Subjects wanted to see more raw weather information incorporated with the tool, as well as navigational information like airspace boundaries.

Question 21 asked subjects if they had any suggestions to make flight planning with the Honeywell prototype easier? (*e.g.,* additional functionality/ information to incorporate, better display formats, tools to help you make decisions with this data, etc.). A summary of comments and observations is as follows:

- Combine *Concept-A* and *Concept-B* (S1, S3, S10, S28).
- Display raw weather information as overlays on flight plan maps. (S2, S7, S12, S13, S17, S19). Better visuals of weather (S18), tailored to specific routes and time period (S26).

- Cleaner more modern interface, *i.e.,* terrain features, airspace overlay (S5).
- 3D holographic type imaging (S6).
- The animation needs work. Make it so you can left click and put the plane on a certain point on the route (S9).
- Increase font size (S12).
- It would be nice if another person could work together at the same time to optimize the planning (S17).
- Adaptation to other airframes (S19).
- More robust toggling (S19).
- The weather info needs to be better organized for quicker reference, and omission of unneeded info (S21).
- Allow the user to view weather data used to create hazard areas (S22, S25).
- State letter on each state in the map (S27).
- A tool that will analyze weather and fuel burn and the best routing solution (S31) [Note: it is possible that the subject did not know about the Route Comparison Tool].

A popular suggestion was combining the features of *Concept-A* (mainly polygons) and *Concept-B* (route modification abilities). Another suggestion that has come up frequently in the questionnaire is the inclusion of raw weather overlays in the tool. The current iteration of the Honeywell prototype planner includes all these features, but for experimental purposes some features were restricted to one *Concept* (polygons vs. route modification) and some features were disabled (raw weather information overlays).

Question 22 asked subjects for their perception of the validity of the experimental equipment, information sources, and scenarios, and efficacy of training. The intent of this question was to verify that the experimental setup, the scenarios, and the weather data used were realistic enough to serve as the foundations of the experiment. A summary of comments and observations is as follows:

- I thought the experiment was quite valid (S1, S2, S5, S11, S12, S19, S25, S30). Realism was fairly accurate (S3, S4, S11, S25, S30).
- The information was sufficient for the scenarios (S1). The training, sources and scenarios were good (S20).
- The equipment was good (S26).
- Good simulation of real flight planning scenarios (S4).
- The experiment was good (S13, S14, S16, S17), very good (S6, S7, S15, S18), great (S27), outstanding (S29) and very effective in every way (S6).
- I was told that the weather was real weather from a certain time, but it seemed a little far fetched at times (S9).
- I thought the MOCK system needed more useable data *i.e.,* not text, more pictures (S12, S26).

- I think more time for each case is needed so that the validity can be increased. Currently, on a scale of 1-10, I would rate the validity of the equipment at 6 (S21).
- Good, if a little bit too much training (S22). I thought it was great. It could use a little more training (S23, S26).
- It might be somewhat close to reality but not close as it could be (S31).

Subjects clearly felt that the experiment was valid and realistic. As a controlled experiment, where the experimental design called for the control of any variable not under test, there was some concern that the resulting protocol would be too unrealistic. While the trial protocol was highly constrained to only route selection within an environment of limited weather information, most subjects felt that the scenarios were realistic, and the weather data (while lacking some desired information) was sufficient.

Question 23 asked subjects if they would want to use this tool? The intent of this question was to simply elicit feedback on subject opinion of usefulness. A summary of comments and observations is as follows:

- Maybe (S1) if:
  - *Concept-A* and B are combined (S1, S10) and more detailed weather information is available (S1, S12).
  - If I know it is 100% reliable (S13).
- Yes (S6, S11, S17, S28) since it:
  - increased overall situation awareness (S2, S8, S29).
  - decreased fuel costs (S2, S9, S16, S18).
  - more information in flight planning (S3, S21, S27), different perspectives (S15).
  - makes flight planning easier (S4, S8, S14, S16, S19, S22, S23, S24, S26) and safer (S4, S7, S9, S18, S24).
  - has the potential to fill in the communication gap between dispatcher and meteorologists (S5).
  - for planning flights over a great distance this appears to be an excellent tool (S20).
  - for *Concept-B*, The planning was easy even though I still had to decode weather products (S25).
- No, the options are not close enough to the real world of flying. The system is simply too simple (S31).

Most subjects (27 of 31) said they would use the tool, with three more saying they would use a modified version of the tool. One subject felt it was not close enough to a real flight planning tool to use.

Question 24 asked subjects if they were overly taxed on any of the TLX workload dimensions? . A summary of comments and observations is as follows:

- Frustration - its ease of use made me wonder if it calculated correctly! (S2)

- Not really, however, not having the backup of a visually predicted weather hazard in *Concept-B* (S7) did require me to think more and work harder (S3, S18), increasing the frustration factor (S1, S3, S26) and mental demand (S3, S4). Performance was not very confident with *Concept-B* (S26).
- Not that much (S18), No (S6, S7, S8, S11, S16, S17, S20, S21, S24, S25, S27, S28, S29), not at all (S12, S19, S30, S31), absolutely not (S5)
- Frustration was high but mainly because of the program use with animation (S9)
- Temporal demand - I thought more time would be needed (S10). It is hard to get a 3D picture using only MOCK (S14).
- At first I was a little frustrated because there was so much. But after a while it was easy to use (S23)
- No, the software helped eliminate these workloads (S15).
- I became frustrated at my lack of knowledge (how to read and understand) weather products provided. I would have liked to create hazard areas (heights, movement etc) (S22)

Most subjects did not feel overly taxed on any of the workload dimensions. Some subjects expressed an increase in frustration when using *Concept-B* since the weather information was not integrated with route information. Two subjects felt increased levels of temporal demand when using MOCK, preferring more time to complete the trials.

6. Discussion

6.1 *General Discussion*

The subject pool contained various levels of experience with flight planning tool and meteorology training. Across subjects, there was strong support for the integration of processed weather information in the form of polygons. The experimental results detailed in the previous section revealed a significant effect of *Concept* in every measurement where the condition was present. The principal benefit of *Concept-A* was the inclusion of weather polygons representing area of hazardous weather. Subjects rated their trust in the polygon definition, boundaries, and severities as high.

Workload of all types was significantly reduced in *Concept-A* over *Concept-B*, where subjects were required to mentally integrate weather and route information across two screens and two applications.

The "Distance Flown in Hazard" metric endeavors to measure dispatchers success level in avoiding hazardous weather, *as defined by the staff meteorologist*. By superimposing the polygon hazards on the selected route for a given subject on a given trial, one can calculate the number of miles that are flown inside a weather polygon (above severity threshold) in both *Concept* conditions. On average, subjects under *Concept-B* flew almost six times as many miles within an area of hazardous weather (as defined by the meteorologist) than subjects under *Concept-A*. More interesting than just looking at averages, however, is to look to see how many subjects in each condition were successful in avoiding the hazardous weather. Of the 192 trials under *Concept-A*, a route that penetrated a hazard was selected 22 times, or approximately 11.5% of the time. Subjects using the tool under the *Concept-B* condition penetrated a hazard in 139 of 192 trials (74.2% of the time). In other words, subjects were over six times as likely to select a route that penetrates areas of weather determined by the meteorologist to be too severe to fly through. It should be noted that there was little difference between *Concept-A* and *Concept-B* in the number of miles once a route that penetrates weather has been chosen (230 miles versus 256 miles, respectively).

On average, across all 12 routes, fuel use in *Concept-B* was 9.0% higher than fuel use under *Concept-A*. Planning time averaged 250 seconds (4:10 min:sec) under *Concept-A*, while *Concept-B* trials averaged 430 seconds (7:10 min:sec).

Several analyses looked to assess whether the experiment and its assumptions were valid. Did subjects look at MOCK while trying to route around weather? Yes, subjects accessed MOCK 2310 times, with a greater share of "hits" in *Concept-B* (1415 of 2310, or 61.2%). Did the weather familiarization session allow subjects to build a general situation awareness of the *Weather Case*? Yes, as over half of the subjects were able to correctly answer very specific, detailed questions after looking at a *Weather Case* for five minutes. Did subjects subjectively feel that the scenarios were valid, that the data used was realistic, that the tasks they were asked to perform were typical of dispatchers? Subject comment demonstrated that they clearly felt that the experiment was valid and realistic.

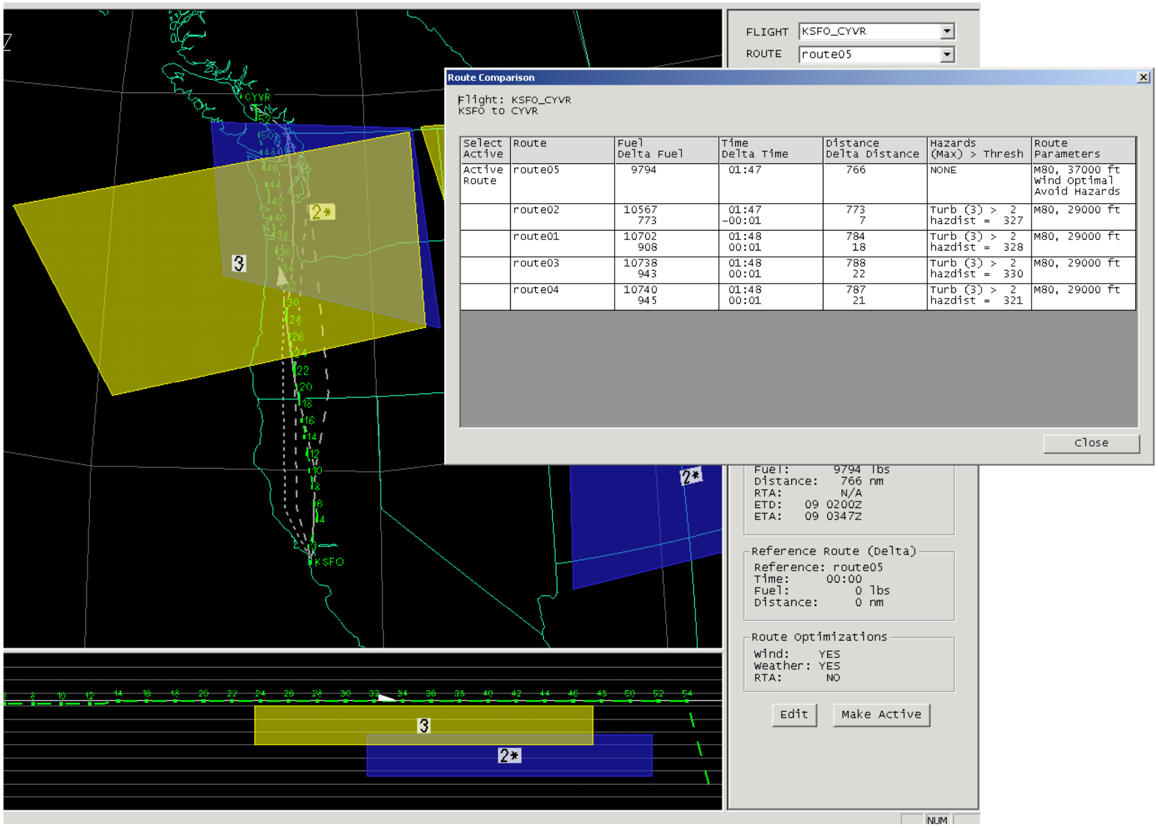
Subjects rated the validity of the experiments as a 5.8 on a seven-point scale. The scenario's validity rested primarily on the fact that real data was used, and the routes were consistent with real operations. The weather came from actual recorded weather, and the company routes were selected from a database of actual routes flown by airlines. Subjects expressed concern that MOCK data, while rated good enough, was missing data they wanted, or at least data in the format they wanted (*e.g.,* AIRMETs).

Subjects also provided formative feedback in the form of comments and answers to questions the questionnaires.

Deeper inspection of the experimental data reveals that there were different strategies that subjects used to accomplish their task of selecting a fuel-efficient route that avoided hazardous weather. The next two sections will look at two routes in detail to understand subject behavior and interaction with the tool.

6.2 *Vertical Solution*

Route 5, illustrated in Figure 16, required a subject to dispatch a flight departing from San Francisco (SFO) to Vancouver (YVR). The weather situation en-route consists of icing and turbulence. The meteorologist characterized the area of icing at severity level 2, and the area of turbulence at severity level 3 (see Appendix D for a definition of severity levels). Thus the route may pass through the icing but not the turbulence, since subjects were allowed to fly through weather of severity level 2 and below. In addition, the icing tops were at 24,000 feet, too low to be of concern. The turbulence topped out at 35,000 feet however, considerably higher than the company route nominal cruising altitude of 29,000 feet. As can be seen in Figure 16, a lateral deviation would be excessive due to the large extent of the area of turbulence, and so a vertical solution would be more efficient. Of the 16 subjects conducting this trial in *Concept-A*, 15 chose the AWIN-optimized route (*Route 5*) with a cruise altitude of 37,000 feet. The other subject chose another route (*Route 2*), at 29,000 feet, and thus penetrated through the center of the hazard for a distance of 327 miles.



Select Active	Route	Fuel Delta Fuel	Time Delta Time	Distance Delta Distance	Hazards (Max) > Thresh	Route Parameters
Active Route	route05	9794	01:47	766	NONE	M80, 37000 ft Wind Optimal Avoid Hazards
	route02	10567 773	01:47 -00:01	773 7	Turb (3) > 2 hazdist = 327	M80, 29000 ft
	route01	10702 908	01:48 00:01	784 18	Turb (3) > 2 hazdist = 328	M80, 29000 ft
	route03	10738 943	01:48 00:01	788 22	Turb (3) > 2 hazdist = 330	M80, 29000 ft
	route04	10740 945	01:48 00:01	787 21	Turb (3) > 2 hazdist = 321	M80, 29000 ft

Figure 16. Route 5 from San Francisco (SFO) to Vancouver (YVR).

The solutions of the subjects dispatching *Route* 5 under *Concept-B* are listed in Table 48. Of the 16 subjects in *Concept-B*, only four attempted a vertical solution, and were the only subjects that solved the problem successfully. The six subjects that modified a company route laterally did not successfully avoid the area of turbulence. Six other subjects chose an existing company route. The 12 subjects who deviated laterally or chose a company route averaged a hazard penetration distance of 328 miles.

Table 48. Characterization of Route 5 solutions for the 16 subjects using *Concept-B*.

Subject # (*Concept-B*)	Modify a Route?	Altitude of Solution	Hazard Penetration Distance
5	**Vertical**	**35,000**	**0**
6	Company route	29,000	327
7	Lateral	29,000	328
8	Company route	29,000	327
9	Lateral	29,000	332
10	Company route	29,000	327
11	Company route	29,000	327
12	**Vertical**	**35,000**	**0**
21	Company route	29,000	327
22	Company route	29,000	327
23	Lateral	29,000	322
24	Lateral	29,000	349
25	**Vertical**	**41,000**	**0**
26	Lateral	29,000	322
27	**Vertical**	**37,000**	**0**
28	Lateral	29,000	324

Route 5 differed from most of the other routes in the workload required to complete this trial. As reported in section 5.2, Mental workload, physical workload, temporal demand, and average workload all showed that *Route* 5 was significantly different from at least one other route. For the subjects that modified the route laterally, most deviation were rather small and involved adding only one or two additional waypoints. A vertical solution only requires the changing of the cruise altitude. Apparently, subjects found the workload associated with this trial to be less than most of the other trials. In *Concept-A*, *Route* 5 also had the lowest amount of planning time, with an average of 145 seconds as compared to an average planning time of 243 seconds for *Routes* 1-6 and 255 seconds for *Routes* 7-12 in *Concept-A*.

6.3 *Lateral Solution*

Subjects reported the highest average workload for *Route* 6, illustrated in Figure 17. A large convection hazard crosses all the company routes, traveling northeast at 60 degrees. A smaller area of convection lies southeast of the destination, but is not moving. In *Concept-A*, only 75% of the subjects chose the route that avoided all weather (route05), which weaves its way between the two areas of convection. This represented the poorest performance of the subjects in *Concept-A* across the 12 routes (see Table 30).

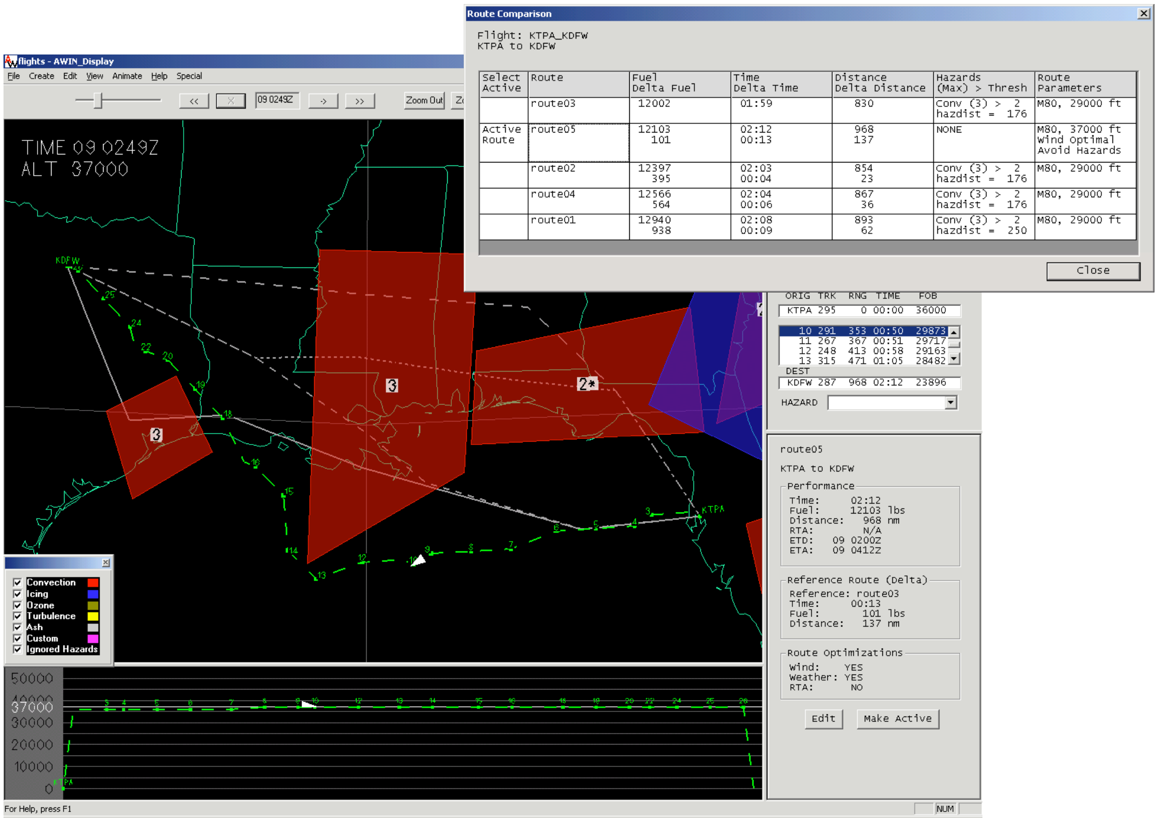


Figure 17. Route 6 in Concept-A. The optimal route is active.

Of the 16 Subjects dispatching *Route* 6 in *Concept-B*, 13 failed to avoid the hazardous weather, as defined by the meteorologist. Subject's solutions are summarized in Table 49. Thirteen subjects modified routes, with six subjects taking a northern deviation, three deviating southwards, two modifying routes slightly but still flying straight trough the area, and one subject who attempted (unsuccessfully) a vertical deviation. None of the company routes avoided the area of hazardous weather. Of the three successful solutions, two modified a route to the north, and one to the south.

Table 49. Subject solutions for Route 6, under *Concept-B*.

Subject # (Concept-B)	Modify a Route?	Hazard Penetration Distance	Fuel Use
5	**Lateral (North)**	**0**	**13,668**
6	Straight Through	216	13,667
7	**Lateral (North)**	**0**	**14,192**
8	Company Route	176	12,556
9	Lateral (South)	33	12,417
10	Straight Through	176	12,271
11	Lateral (North)	97	12,997
12	Lateral (North)	178	12,545
21	Company Route	176	12,002
22	Company Route	176	12,002
23	Lateral (North)	143	13,166
24	Lateral (North)	180	13,492
25	Lateral (South)	221	11,940
26	Lateral (South)	27	14,993
27	**Lateral (South)**	**0**	**15,766**
28	Vertical (37,000 ft)	177	11973

The three successful solutions are characterized by the large amount of fuel expended, averaging 14,542 pounds. This is due to the wide berths that the subjects gave the area of weather, as illustrated by Subject 7's and 27's solution shown in Figure 18. The solutions are shown at the point in the flight of closest approach to the weather hazard. Of course, subjects did not see the weather polygons depicted on Figure 18 since they were in *Concept-B*. Either the subjects had a different conception of the geographical extent of the convection en-route, or their solution strategy involved a large deviation (or buffer zone) to avoid possible conflicts with hazardous weather. In contrast, the AWIN-optimized route required 12,103 pounds of fuel, principally because it was able to weave its way through the weather rather than taking large deviation around it.

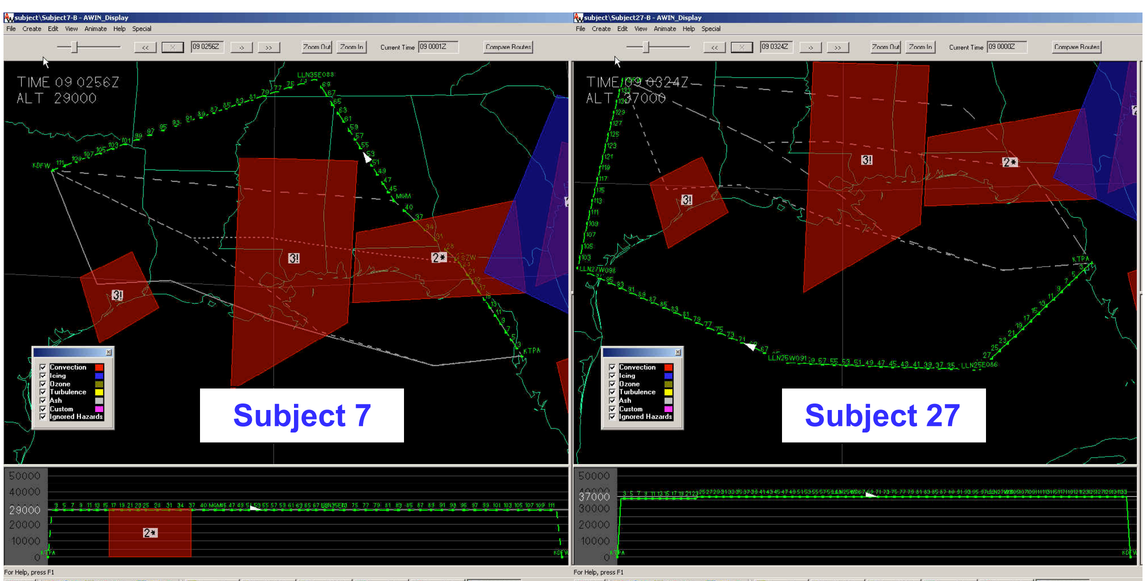


Figure 18. Successful solution strategies of two subjects in Concept-B for Route 6.

Some subjects did attempt to modify a company route and fly through the areas of weather. Subject 9 was almost successful, as illustrated in Figure 19. The flight is shown at two times: (1) two hours before the flight, when the subject is planning the route, and (2) one hour into the flight when the flight flies through a small portion (33 miles) of the area defined by the meteorologist to be hazardous convection. Again, the subject did not have access to the polygons when constructing his solution, and so the subject's solution represent his or her interpretation of the raw weather data, which in this case differed only slightly from that of the meteorologist.

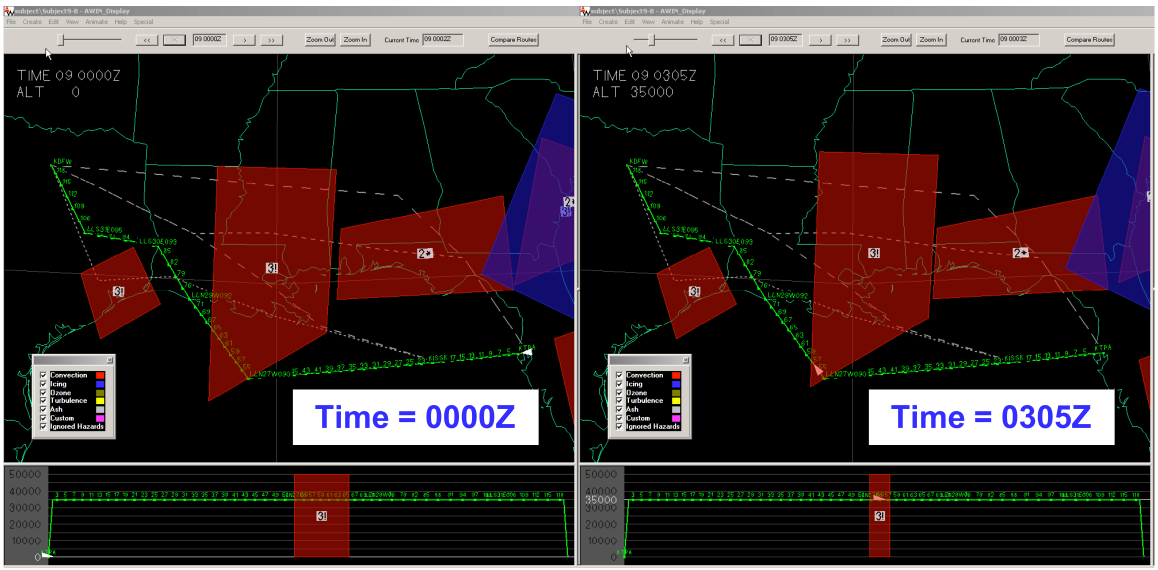


Figure 19. Subject 9 solution for Route 6 under *Concept-B*.

Of the 12 routes, subjects in *Concept-B* spent the most amount of time planning *Route* 6. The *Route* 6 average planning time was 542 seconds, compared to an average of 457 seconds for *Routes* 1-6 and 402 for *Routes* 7-12 in *Concept-B*.

6.4 *Related Work in Weather Products*

This experiment was undertaken to assess the potential viability of the Honeywell Prototype flight planning system. Of particular interest was an empirical evaluation of the integration of processed weather information with route information. The results of the experiment indicate that the approach has promise. One aspect of the system is the rating of polygon severity on a linear scale. Work in the weather products industries continues, with several new developments that are relative to the Honeywell effort in this area. At a recent meeting of the SC 195 Flight Information Services committee (May 29-30, 2002), The National Center for Atmospheric Research (NCAR) provided a status update on NCAR's various weather forecast products (see http://adds.aviationweather.noaa.gov/ for additional information on latest developments). The three main products under development by NCAR that have particular relevance to the AWIN decision support tool are for Turbulence, Icing, and Convection. Currently the scales for convection are 1-6 with 6 indicating the highest intensity. This scale is consistent with previous information received from NCAR and consistent with the AWIN model. NCAR is using polygons to describe the weather in hourly forecasts. While some of the NCAR polygons are concave, the data can be manipulated to be consistent with the conceptual input AWIN weather model.

3. Acknowledgements

This work was supported in part by NASA Langley Research Center Cooperative Agreement NCC-1-362 under the technical supervision of Dr. Kara Latorella. The authors would like to acknowledge the contributions of Dr. Chris Herbster, Mike Masscotte, and Ian Wilson of Embry-Riddle Aeronautical University, for the weather data capture / processing, subject recruitment, company route database, use of experimental facilities, and other contributions. The authors would also like to acknowledge Thea Feyereisen, Gary Hartmann, Chris Misiak, Dr. Vic Riley, Robert L. Schultz, Don Shaner, Jennifer Sly, and Fred Wagner, who have all significantly contributed to the development of the AWIN tool. We would like to thank Dr. Trish Ververs and Dr. Bill Rogers for their help in the experimental analysis. We would like to thank Subramani Krishnamurthi for assistance processing the data. Finally, we would like to thank all the participants of the dry run, for the detailed feedback that greatly improved the final experiment plan, and the experiment subjects, for their conscientious efforts.

4. References

[1] *Airline Operational Control Overview,* Prepared by the Airline Dispatchers Federation and Seagull Technology, Inc., Under the direction of the Federal Aviation Administration and Boeing Commercial Airplane Group per Cooperative Research and Development Agreement 93-CRDA-0034, May 1995.

[2] Box, G.E.P., Hunter, W.G., and Hunter, J.S. (1978). *Statistics for Experimenters*, John Wiley & Sons, New York.

[3] COMET Case Studies (online). "Frequently Asked Questions about Case Studies on the Web." Available at http://www.joss.ucar.edu/cometCases/faq.html.

[4] GRIB Table B - Grid Identification (online). Available at http://www.nco.ncep.noaa.gov/pmb/docs/on388/tableb.html

[5] JOSS Web Site (online). Available at http://www.joss.ucar.edu/.

[6] Dorneich, M. C., Pratt, Steve, Olofinboba, O., and Osborne, D. (2002). "AWIN AWIN Weather Avoidance using Route Optimization as a Decision Aid: Weather Survey Source Results, Revised Experiment Plan, and Revised Integration Plan." *NASA/CR 2002*. Prepared for NASA Langley Research Center under contract NCC-1-362, Milestone #5 Technical Report, Honeywell Laboratories, March 15, 2002.

[7] Olofinboba, Olu, Dorneich, Michael, Osborne, Dannielle, and Ververs, Patricia, (2001). "AWIN Weather Avoidance using Route Optimization as a Decision Aid: An AWIN Topical Study. Experimental Plan", prepared for NASA Langley Research Center under contract NCC-1-362, Milestone #4 Technical Report, *Honeywell Laboratories*, September 29, 2001.

[8] *Pilotbrief on the Net.* WSI Corporation Web Site, http://www.wsicorp.com/wsicorp/markets/markets_media_a_gal_H.htm, Sept 2001.

[9] Pratt, Steve, Meisner, John, Wagner, Fred, Shaner, Donald, and Hartman, Gary (2001). "Weather Avoidance Using Route Optimization as a Decision Aid: An AWIN Topical Study: Specific Enhancements and Integration Test Plan", prepared for NASA Langley Research Center under contract NCC-1-362, Milestone #4 Report, *Honeywell Laboratories*, September 2001.

[10] *TWA Flight Dispatch Training Manual,* Trans World Airlines, 2000.

[11] *Weather Avoidance Using Route Optimization as a Decision Aid: An AWIN Topical Study*, Phase 1 Report, Honeywell Technology Center, Minneapolis, MN, December 30, 1998.

[12] *Weather Avoidance Using Route Optimization as a Decision Aid: An AWIN Topical Study*, Final Report, Honeywell Technology Center, Minneapolis, MN, November, 1999.

[13] WSI home page (online). Available at http://www.wsi.com/sites/sites.htm. 2002.

[14] Winer, B.J. (1971). *Statistical Principles in Experimental Design*. New York: McGraw Hill Book Company.

Appendix A. Counterbalanced Order of Trials

Table 50 shows the randomized order of trials, generated from a computer program.

Table 50. Experimental Worksheet.

	Standard-AWIN trials						Modified-AWIN trials					
Trial	1	2	3	4	5	6	7	8	9	10	11	12
Dispatcher 1	W1(R3)	W1(R1)	W1(R2)	W2(R5)	W2(R6)	W2(R4)	W4(R11)	W4(R10)	W4(R12)	W3(R7)	W3(R8)	W3(R9)
Dispatcher 2	W2(R6)	W2(R5)	W2(R4)	W1(R1)	W1(R3)	W1(R2)	W3(R8)	W3(R7)	W3(R9)	W4(R11)	W4(R10)	W4(R12)
Dispatcher 3	W1(R2)	W1(R3)	W1(R1)	W2(R6)	W2(R4)	W2(R5)	W3(R9)	W3(R8)	W3(R7)	W4(R12)	W4(R11)	W4(R10)
Dispatcher 4	W1(R1)	W1(R3)	W1(R2)	W2(R6)	W2(R4)	W2(R5)	W3(R9)	W3(R7)	W3(R8)	W4(R11)	W4(R12)	W4(R10)
Dispatcher 9	W4(R11)	W4(R12)	W4(R10)	W3(R9)	W3(R8)	W3(R7)	W2(R5)	W2(R4)	W2(R6)	W1(R3)	W1(R2)	W1(R1)
Dispatcher 10	W4(R10)	W4(R11)	W4(R12)	W3(R8)	W3(R9)	W3(R7)	W1(R2)	W1(R1)	W1(R3)	W2(R4)	W2(R6)	W2(R5)
Dispatcher 11	W3(R7)	W3(R8)	W3(R9)	W4(R12)	W4(R10)	W4(R11)	W2(R4)	W2(R6)	W2(R5)	W1(R2)	W1(R3)	W1(R1)
Dispatcher 12	W3(R7)	W3(R9)	W3(R8)	W4(R12)	W4(R10)	W4(R11)	W2(R5)	W2(R6)	W2(R4)	W1(R3)	W1(R1)	W1(R2)
Dispatcher 17	W2(R6)	W2(R4)	W2(R5)	W1(R2)	W1(R3)	W1(R1)	W4(R11)	W4(R12)	W4(R10)	W3(R9)	W3(R7)	W3(R8)
Dispatcher 18	W1(R3)	W1(R1)	W1(R2)	W2(R4)	W2(R5)	W2(R6)	W4(R11)	W4(R12)	W4(R10)	W3(R7)	W3(R9)	W3(R8)
Dispatcher 19	W1(R3)	W1(R2)	W1(R1)	W2(R4)	W2(R6)	W2(R5)	W3(R7)	W3(R8)	W3(R9)	W4(R11)	W4(R10)	W4(R12)
Dispatcher 20	W1(R3)	W1(R2)	W1(R1)	W2(R5)	W2(R6)	W2(R4)	W4(R12)	W4(R11)	W4(R10)	W3(R8)	W3(R9)	W3(R7)
Dispatcher 25	W3(R7)	W3(R9)	W3(R8)	W4(R11)	W4(R12)	W4(R10)	W1(R2)	W1(R3)	W1(R1)	W2(R4)	W2(R5)	W2(R6)
Dispatcher 26	W4(R11)	W4(R10)	W4(R12)	W3(R9)	W3(R7)	W3(R8)	W2(R4)	W2(R6)	W2(R5)	W1(R1)	W1(R2)	W1(R3)
Dispatcher 27	W4(R12)	W4(R11)	W4(R10)	W3(R9)	W3(R8)	W3(R7)	W1(R1)	W1(R3)	W1(R2)	W2(R5)	W2(R6)	W2(R4)
Dispatcher 28	W4(R12)	W4(R11)	W4(R10)	W3(R7)	W3(R9)	W3(R8)	W2(R6)	W2(R4)	W2(R5)	W1(R1)	W1(R2)	W1(R3)

	Modified-AWIN trials						Standard-AWIN trials					
Trial	1	2	3	4	5	6	7	8	9	10	11	12
Dispatcher 5	W2(R4)	W2(R5)	W2(R6)	W1(R2)	W1(R1)	W1(R3)	W4(R12)	W4(R11)	W4(R10)	W3(R9)	W3(R7)	W3(R8)
Dispatcher 6	W2(R5)	W2(R4)	W2(R6)	W1(R2)	W1(R1)	W1(R3)	W4(R11)	W4(R10)	W4(R12)	W3(R7)	W3(R9)	W3(R8)
Dispatcher 7	W1(R3)	W1(R2)	W1(R1)	W2(R5)	W2(R4)	W2(R6)	W3(R8)	W3(R9)	W3(R7)	W4(R11)	W4(R10)	W4(R12)
Dispatcher 8	W2(R4)	W2(R6)	W2(R5)	W1(R3)	W1(R1)	W1(R2)	W3(R9)	W3(R8)	W3(R7)	W4(R12)	W4(R10)	W4(R11)
Dispatcher 13	W3(R7)	W3(R9)	W3(R8)	W4(R10)	W4(R11)	W4(R12)	W1(R1)	W1(R3)	W1(R2)	W2(R5)	W2(R4)	W2(R6)
Dispatcher 14	W4(R10)	W4(R12)	W4(R11)	W3(R7)	W3(R9)	W3(R8)	W1(R1)	W1(R3)	W1(R2)	W2(R4)	W2(R5)	W2(R6)
Dispatcher 15	W4(R10)	W4(R11)	W4(R12)	W3(R9)	W3(R8)	W3(R7)	W1(R2)	W1(R3)	W1(R1)	W2(R5)	W2(R4)	W2(R6)
Dispatcher 16	W3(R8)	W3(R7)	W3(R9)	W4(R11)	W4(R12)	W4(R10)	W2(R5)	W2(R4)	W2(R6)	W1(R3)	W1(R1)	W1(R2)
Dispatcher 21	W1(R3)	W1(R1)	W1(R2)	W2(R6)	W2(R4)	W2(R5)	W4(R12)	W4(R10)	W4(R11)	W3(R9)	W3(R7)	W3(R8)
Dispatcher 22	W2(R4)	W2(R6)	W2(R5)	W1(R2)	W1(R1)	W1(R3)	W4(R11)	W4(R12)	W4(R10)	W3(R8)	W3(R9)	W3(R7)
Dispatcher 23	W1(R1)	W1(R2)	W1(R3)	W2(R6)	W2(R5)	W2(R4)	W3(R9)	W3(R8)	W3(R7)	W4(R10)	W4(R12)	W4(R11)
Dispatcher 24	W2(R5)	W2(R6)	W2(R4)	W1(R3)	W1(R2)	W1(R1)	W3(R7)	W3(R9)	W3(R8)	W4(R12)	W4(R11)	W4(R10)
Dispatcher 29	W4(R10)	W4(R12)	W4(R11)	W3(R9)	W3(R7)	W3(R8)	W2(R4)	W2(R5)	W2(R6)	W1(R1)	W1(R3)	W1(R2)
Dispatcher 30	W3(R9)	W3(R7)	W3(R8)	W4(R10)	W4(R12)	W4(R11)	W1(R2)	W1(R1)	W1(R3)	W2(R5)	W2(R4)	W2(R6)
Dispatcher 31	W4(R12)	W4(R10)	W4(R11)	W3(R8)	W3(R7)	W3(R9)	W1(R3)	W1(R1)	W1(R2)	W2(R6)	W2(R5)	W2(R4)
Dispatcher 32	W4(R11)	W4(R10)	W4(R12)	W3(R8)	W3(R7)	W3(R9)	W1(R3)	W1(R2)	W1(R1)	W2(R4)	W2(R5)	W2(R6)

Appendix B. Derivation of the Experimental Analysis

The Experimental Design of section 2.1.3 can be re-organized, as shown in Table 51. Due to the constraints listed in section 4.1, it was not possible to do a pure within-subjects design (where every subject sees every combination of variables) and thus the experiment was complicated by the nested effect of Wx Case and Route. If the data is split into two groups, one group for *Route* 1-6 and the second group for routes 7-12, then a straightforward analysis is possible for each group of data.

Table 51. Experimental data collected as for each combination of independent variables.

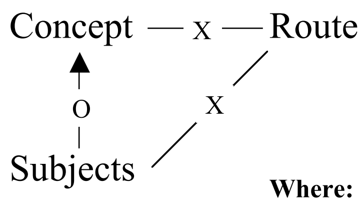
Route	Concept A				Concept B			
	Wx 1	Wx 2	Wx 3	Wx 4	Wx 1	Wx 2	Wx 3	Wx 4
1	G1				G2			
2	G1				G2			
3	G1				G2			
4		G1				G2		
5		G1				G2		
6		G1				G2		
7			G2				G1	
8			G2				G1	
9			G2				G1	
10				G2				G1
11				G2				G1
12				G2				G1

The resulting ANOVA design is a mixed design with *Concept* as a between-subjects variable, and *Route* and *Wx Case* as within-subject variables (having *Route* effects nested under the *Wx Case* variable). The derivation of the analysis follows as such. Section B.1 will present the analysis of a simple 2 x 6 mixed design. Section B.2 will present the analysis of a 2(3) nested design. The notation and derivation is mirrored with similar examples found in [14]. Finally section B.3 will present the combination that results in a 2 x 2(3) design.

B.1 Mixed Design

If we first consider a 2 x 6 mixed design (ignoring for the moment that Route is nested within Wx Case). The experimental design is depicted in Figure 20.

Experimental Design



Subject pool

	R_1	R_2	R_3	R_4	R_5	R_6
C_1	G_1	G_1	G_1	G_1	G_1	G_1
C_2	G_2	G_2	G_2	G_2	G_2	G_2

Where:
G_i a unique group of subjects
c=2 number of concepts
r=6 number of routes
n=16 number of subjects in group G_i

Figure 20. A 2 (Concept) x 6 (Route) with Repeated Measures Experiment.

Such an experiment would have an ANOVA design shown in

Table 52.

Table 52. ANOVA table for a 2 x 6 mixed design.

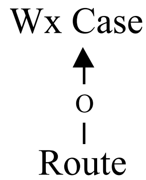
Source of Variation		DOF Formula	DOF
Between Subjects		nc-1	31
C	Concept	c-1	1
Subjects (Concept) [pooled]		c(n-1)	30
Within Subjects		nc(r-1)	160
R	Route	r-1	5
C x R	Interaction	(c-1)(r-1)	5
R x Subjects (Concept)		c(n-1)(r-1)	150

B.2 Nested Designs

When we consider that Route is nested within Wx Case, however, we must revise the design of Figure 20.

Let us first consider a nested design without the Concept between-subjects factor, we would have the ANOVA in Figure 21.

Experimental Design **Subject pool**



W1			W2		
R_1	R_2	R_3	R_4	R_5	R_6
n	*n*	*n*	*n*	*n*	*n*

Where:
w=2 number of weather cases
r=3 number of routes per weather case
n=16 number of measurements

Figure 21. Nested 2(3) Wx Case (Route) design.

The difference between mean effect of *Wx Case* 1 and the mean effect of *Wx Case* 2 will be due in part to differences between the unique effects associated with *Routes* 1-3 and the unique effects associated with *Routes* 4-6.

The unique effects associated with *Routes* 1-3 are confined to *Wx Case* 1 whereas the unique effects associated with *Routes* 4-6 are confined to *Wx Case* 2. The *Route* effects are nested under the *Wx Case* effect. There is no way of evaluating the interaction between the *Routes* and the *Wx Case*. The ANOVA table is given in Table 53.

Table 53. ANOVA table for a 2(3) nested design.

Source of Variability		DOF Formula	DOF
W	*Wx Case*	w-1	1
R (w. a_1)	*Routes* within *Wx Case* 1	r-1	2
R (w. a_2)	*Routes* within *Wx Case* 2	r-1	2
Error	Within *Routes*	wr(n-1)	90

The expected values of the mean squares in this analysis are given in Table 54.

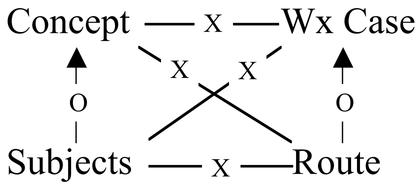
Table 54. ANOVA table for a 2(3) nested design where Route effects have been pooled.

Source of Variability		DOF Formula	DOF
W	*Wx Case*	w-1	1
R (pooled)	*Routes*	w(r-1)	4
Error	Within cell	wr(n-1)	90

B.3 Mixed Design with Nested Design

If we combine the lessons learned from the previous two analysis, we can derive the ANOVA design of a 2 (*Concept*) x 2(3) *Wx Case* (*Route*) experiment. Because of the nesting of subjects within *Concept* and the nesting of *Routes* within *Wx Case*, there can be no three-way interaction terms.

Experimental Design



Subject pool

	W_1			W_2		
	R_1	R_2	R_3	R_4	R_5	R_6
C_1	G_1	G_1	G_1	G_1	G_1	G_1
C_2	G_2	G_2	G_2	G_2	G_2	G_2

Where:
G_i a unique group of subjects
c=2 number of concepts
w=2 number of weather cases
r=3 number of routes
n=16 number of subjects in group G_i

Figure 22. A 2 (Concept) x 2(3) Wx Case(Route) with Repeated Measures Design.

The ANOVA table for the design in Figure 22 is given in Table 55.

Table 55. ANOVA table for a 2x2(3) with repeated measures.

Source of Variability		DOF Formula	DOF
C	*Concept*	c-1	1
W	*Wx Case*	w-1	1
R (W)	*Routes* w/i *Wx Case* (pooled)	w(r-1)	4
Subjects (C)	*Subjects* w/i *Concept* (pooled)	c(n-1)	30
C x W	Interaction	(c-1)(w-1)	1
Subject(C) x W	Interaction	c(n-1)(w-1)	30
R(W) x C	Interaction	w(r-1)(c-1)	4
Subject(C) x R(W)	Error	cw(n-1)(r-1)	120

Appendix C. Experiment Protocol

This appendix contains the experimenter instructions detailing how to run the experiment.

C.1 Experiment Setup

Before the start of the experiment, the AWIN tool and its associated files must be copied for each subject.

1. On the desktop should be two shortcuts to the executables for *Concept-A* and *Concept-B*.
2. Click on the appropriate shortcut.
3. Click the mouse in the Map View; this will bring up a login dialog box.
4. Enter the login name for subject number XX as "SubjectXX" (Do not use spaces in the login name).
5. The appropriate files will be copied and all log / route data can be found in /subjects/SubjectXX-A (if in Concept A) or /subjects/SubjectXX-B (if in *Concept-B*).

Please refer to counterbalanced trial sheet to determine which experimental booklets to pass out. The code number is on the cover in the upper left-hand corner

C.2 Ground Rules

Ground rules to be read to subjects at beginning of experiment

1. Please no talking to each other during the experiment.
2. Please do not talk to anyone after you leave about the experiment. We want all subjects to start with the same amount of knowledge about the experiment (*i.e.,* none!)
3. Do not look ahead in the booklet. We use the booklet to pace the experiment. The experimenter will tell you when to turn to the next page.
4. You can, however, always look back in your booklet to review training or instructions.
5. The experimenter will read directly from the booklet in order to endure that all subjects hear the same briefing.
6. Please feel free to ask questions at any time.

C.3 Experiment Protocol

Subjects are given a numbered experiment briefing guide (see Appendix D). Each Experiment Guide contains all the materials (*e.g.,* consent forms, training briefs, experimental trials) including the correct order of trials for the given subject number (see Appendix A for trial order indexed by subject number). The experimenter will largely read from the booklet in order to ensure that each subject is briefed in the same way.

Table 56 describes instructions that tell the experimenter how to proceed through each section of the Experiment Guide.

Table 56. Experimental protocol instructions.

Section	**Protocol**
Ground Rules	Experimenter to review ground rules (see above)
Introduction	Experimenter reads it to subjects.
Purpose of Assessment	Experimenter reads it to subjects.
Evaluation Personnel	Refer to it and move on.
Experiment Schedule	Experimenter summarizes
Consent Form	Experimenter reads top half. Subjects read consent form, sign it. Experimenter co-signs.
Demographics Form	Subjects read over and fill out
NASA TLX Workload Scale Instructions	Experimenter reads it to subjects. • Remind students that they can refer back to directions/training at any time. • Remind subjects of ratings endpoints
General Instructions	Experimenter reads it to subjects
Weather Information Sources	Experimenter reads it to subjects. Experimenter reads "General" column of the weather severity table, allows time for students to read examples for weather types.
AWIN Baseline Training	Experimenter reads introduction paragraph, summarizes remaining sections. Accompanied by live demo via computer projector. Emphasize winds and selecting altitudes
Break	10 minutes. Skip if running behind.
Concept A or B Training	Depending on which block is done first. Experimenter reads it to subjects.
Practice Trial	Ignore MOCK, concentrate on becoming familiar with AWIN. • If B, tell them to modify a route to avoid Tennessee.
First Block of Trials	Experimenter verbally reminds them • of the task, • they have approximately 1 hour (A) or 1 _ hour (B) to complete the 6 trials, and • are to fill out the TLX survey after each trial. Then experimenter remains silent, except to answer questions (only give clarification info)
Complete post-block questionnaire	Subjects fill out.
Break	10 minutes
Concept A or B Training	Depending on which block is done second. Experimenter reads introduction, summarizes remaining sections
Practice Trial	Students can ignore MOCK.
Second Block of Trials	Experimenter verbally reminds them • of the task (Read through scenario directions), • they have approximately 1 hour (A) or 1 _ hour (B) to complete the 6 trials, and • are to fill out the TLX survey after each trial. Then experimenter remains silent, except to answer questions (only give clarification info)
Complete post-block questionnaire	Subjects fill out.
Post-experiment questionnaire	Subjects fill out. • Remind them what Concept A and B were.
Group discussion and debrief	If time permits (< 4 hours). including answering questions
Conclusion	Thank them for their time, tell how to be paid. Remind to not discuss.

Appendix D. Experiment Briefing Materials

This Appendix contains samples of the materials presented to subjects during the experiment.

D.1 Schedule

The experiment schedule is described in Table 57.

Table 57. Expected experiment schedule.

Section	Estimated Duration
Overview of the Experiment • Introduction • Purpose of Assessment • Evaluation Personnel • Experiment Schedule • Consent Form • Demographics Form	20 minutes
Training • NASA TLX Workload Scale Instructions • General Instructions • Weather Information Sources • Baseline AWIN Tool	20 minutes
Training: *CONCEPT-A* Flight Planning Tool • Displays • Tasks • Practice Trial	15 minutes
Break	5 minutes
Block of Experimental Trials: *Concept-A* • Running through the experimental scenarios • Completing the NASA-TLX workload scale	60 minutes
Completing the post-block questionnaires	5 minutes
Break	10 minutes
Training: *Concept-B* Flight Planning Tool • Displays • Tasks • Practice Trial	10 minutes
Block of Experimental Trials: *Concept-B* • Running through the experimental scenarios • Completing the NASA-TLX workload scale	60 minutes
Completing the post-block questionnaires	5 minutes
Final Debrief • Post-Experiment Questionnaire • Question and Answer period	30 minutes

D.2 Consent Form

You are about to participate in a flight planning study co-sponsored by Honeywell and NASA, and run by Embry-Riddle Aeronautical University. The duration of the experiment is not expected to exceed 4 hours. The purpose of this experiment is to investigate user interactions with a tool for optimization and planning that assists with strategic planning and re-planning through single or multiple hazard areas. It will evaluate display and interaction methods for flight planning and flight plan optimization around or through hazards or volumes of airspace considered non-desirable for trajectory penetration, namely severe weather and other hazard areas. You will be interacting with a computer workstation to complete the flight planning tasks.

There will be an experimenter present throughout the experiment. In addition, performance and subjective data will be collected as you perform the tasks in the experiment. Performance will be videotaped with audio for use in connection with this investigation. Data resulting from participation will be held confidentially by the experimenters, will be referenced only by participant number, and summarized to assure anonymity.

Statement of Consent

Thank you for participating in this evaluation. This experiment is considered Honeywell and NASA confidential and proprietary and by signing this sheet, you agree not to divulge any specifics of the experiment, including but not limited to displays and interaction methodology.

I have read and understand the objectives of the experiment and willingly agree to participate. I understand that at anytime I am free to withdraw from the experiment. I agree to treat this evaluation as Honeywell and NASA Confidential and Proprietary.

Participant Signature	Print Participant Name	Date
Experimenter Signature	Print Experimenter Name	Date

Thank you for your participation in this study!

D.3 Dispatcher Demographics

The information contained in this questionnaire will help us understand the experiment results in terms of dispatcher characteristics. All information contained herein will be kept confidential.

GENERAL

Name:	______________________________
Address:	______________________________
Email:	______________________________
Phone: ____________________	Age: ________
Years of education (*e.g.,* B.S. = 16): ________	Gender: M / F

EXPERIENCE

Please describe any experience you have with pre-flight route selection. Include a description of any tool that you may have used.

Please describe any experience you have with meteorology. Include a description of any meteorological tools that you may have used.

Thank you for completing this questionnaire. This information will be used to help us understand results in terms of dispatcher characteristics. All information contained herein will be kept confidential. If you have any additional information you think would be useful, please feel free to write it below.

D.4 NASA-TLX Workload Scale

During the initial training stage, subjects were trained on how to complete NASA Task Load Index (NASA-TLX) (see Section D.4.1). Once trained, subjects filled out a TLX Index (see Section D.4.2 for a sample) immediately after each of the 12 trials.

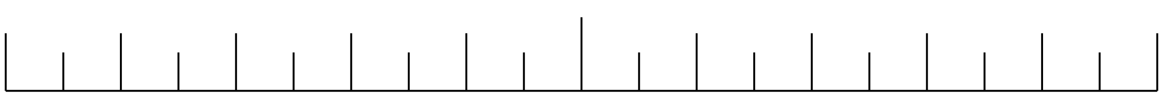
D.4.1 Instructions

There are two ways to assess the effectiveness of decision support tools and their related tasks. One method is to gauge performance using accuracy scores and response time measures. The second method is to consider the perceived level of workload. Your performance will be measured using both methods throughout the experiment. In order to measure your subjective workload, you will be asked to fill out a NASA Task Load Index (NASA-TLX) at the end of all trials.

The section of the NASA-TLX that you will complete consists of six rating scales. Each scale represents an individual workload descriptor: mental demand, physical demand, temporal demand, performance, effort, and frustration. **Place an 'X' along each of the six scales indicating the place along the index that best describes your workload for the trial immediately preceding the administration of the rating scales.** Be sure to note the descriptions associated with each of the scales. Performance has "good" on the left and "poor" on the right, while the rest of the scales have "low" and high" as endpoints. Accompanying the ratings scales is a description of each of the measures. Read the descriptions in order to familiarize yourself with the meanings of the workload descriptors.

- **Mental Demand** – how much mental effort is required to perform the task (*e.g.,* thinking, deciding, remembering)
- **Physical Demand** – how much physical effort is required to perform the task (*e.g.,* pushing, pulling, reaching, stretching)
- **Temporal Demand** – how much time pressure you feel to complete the task (*e.g.,* relaxed pace or fast and furious?)
- **Performance –** how successful you feel you are in completing the task
- **Effort** – how hard you work to complete the task
- **Frustration** – how aggravated or annoyed versus secure or content you feel about accomplishing the task.

Example: **Mental Demand**



Low High

NOTE: When completing the NASA-TLX rating sheet, consider *only* the immediately preceding scenario. Specifically concentrate on the level of workload you experienced in completing the tasks involved with <u>selecting a fuel-efficient route while avoiding hazardous weather</u> in the previous scenario.

D.4.2 Sample NASA-TLX Workload Questionnaire

Instructions: Place an 'X' along each of the six scales indicating the place along the index that best describes your workload *only* for the trial immediately preceding the administration of the rating scales. Specifically concentrate on the level of workload you experienced in while selecting a fuel-efficient route while avoiding hazardous weather. For a description of the six rating scales, please review the description on page 89.

Mental Demand

Low High

Physical Demand

Low High

Temporal Demand

Low High

Performance

Good Poor

Effort

Low High

Frustration

Low High

D.5 General Instructions Given to Subject

You, as the dispatcher, will be given a total of 12 flights to dispatch, in two blocks of six flights, one block in the "*Concept-A* condition", and one block in the "*Concept-B* condition". The two conditions each use a different version of the Honeywell prototype AWIN flight planning system.

In both blocks you will have access to weather information sources normally found in a flight dispatch environment (*e.g.,* convective weather plots, turbulence plots, wind charts, etc). You will be briefed on what weather types and severity are considered hazardous.

For each of the two block of trials, you will be given six city-pairs, one at a time.

Before each set of three flights, you will spend five minutes of dedicated time becoming familiar with the weather (weather familiarization session). After this dedicated time, your will conduct three trials, one after another. The weather is valid for all three trials, and you will continue to have access to all the weather information you had during weather familiarization session. After the third trial, you will spend five minutes of dedicated time becoming familiar with a new set of weather data, which will be applicable to the next three trials. The second block of six trials follows the same pattern where the first three trials will occur with one weather data set, and the second three trials will occur with another weather data set.

It is your job to decide on a fuel-efficient route for the flight that flies between the city pair, and avoids hazardous weather. Each city-pair has an associated set of "company routes". A company route is a route pre-approved by the airline. These routes were created with no consideration of weather. The flight planner in each experimental condition is capable of assessing the predicted fuel usage of each route, based on the day's wind and temperature information. This experiment assumes motivations for route selection are for choosing best route as defined by weather avoidance and fuel efficiency. Other constraints that may be motivating performance in actual operations are not relevant in these scenarios.

In the *Concept-A* condition, an additional route is provided that is optimized for fuel and avoids hazardous weather.

In the *Concept-B* condition, it is possible to modify a company route in order to avoid hazardous weather.

It is your job to choose the most fuel-efficient route that avoids any hazardous weather. You may choose any existing routes, or may modify an existing route in the *Concept-B* condition.

D.6 Training: Weather Information Sources

D.6.1 MOCK Weather Information Sources

For all city-pair scenarios, you will be provided with a set of weather information graphics, accessible through a web page, illustrated in Figure 23. Weather data for the trial was gathered from actual weather events.

There are four set of weather data, each from a different date. Each weather data set is valid for a set of three experimental trials. The left-side column of MOCK provides links to each date, and the right-side main window displays links to all the weather data available for that date.

Before each set of three trials, you will spend five minutes of dedicated time becoming familiar with the weather (weather familiarization session) via the graphics on the MOCK display. You are **not** required to look at every link. The goal of the familiarization period is to allow you a dedicated amount of time to browse the data (weather graphics) and gain a general appreciation of the weather situation. Once the dedicated weather familiarization session is over, you will begin three trials where that weather data is valid. You still have access to the weather data during the trials.

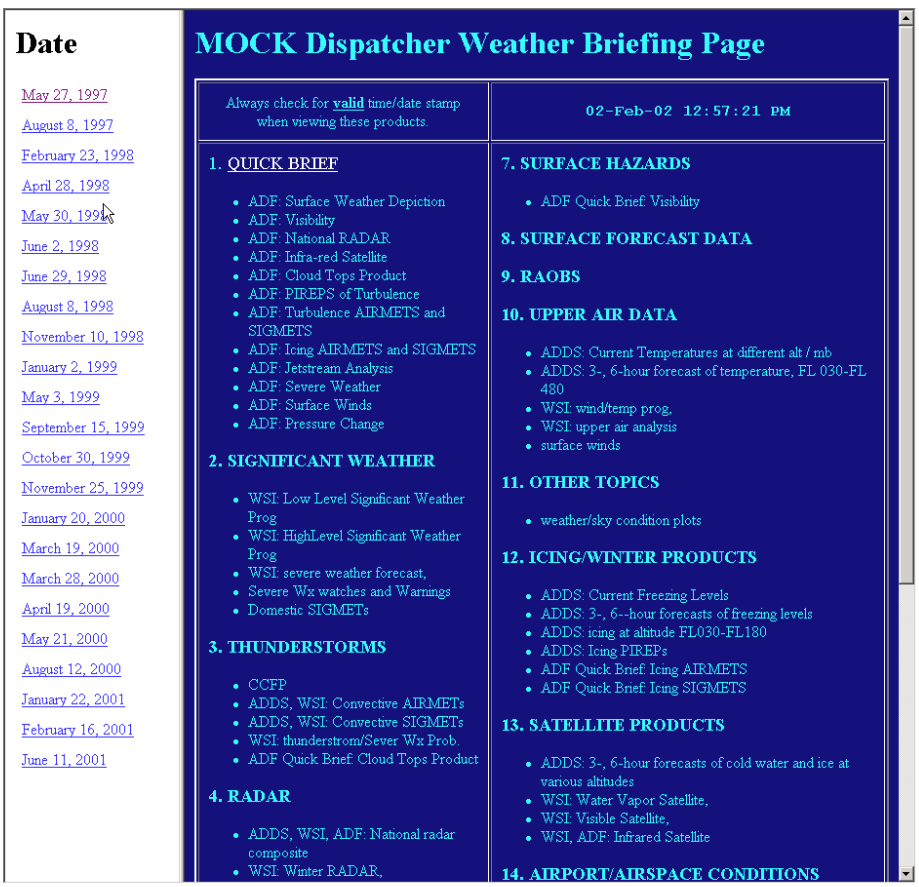


Figure 23. MOCK Dispatcher Weather Information Briefing Page.

D.6.2 Rules for Assessing Weather Boundaries and Severities

During each trial, you will be asked to assess the weather to decide if it impacts your planned route between a city-pair. This section describes the rules you can use to decide if weather severity is significant enough to cause a reroute around the weather phenomena.

In all trials, you will have access to raw weather data via MOCK. You must assess if weather is significantly severe to cause you to reroute around it. Table 58 describes four severity levels, in general and for specific weather types (convection, turbulence, and icing). Examples are given to help calibrate your assessment of weather severity for each type of weather.

For all weather types, routes should not pass through weather of severity level 3 or severity level 4.

Table 58. Description of weather classification scheme per weather type.

Lvl	General	Convective	Turbulence	Icing
1	Weak or poorly organized system exists, low to moderate potential for development	Current RADAR reflectivities in the 20-30 DBz range (or lower)	Modest wind shear (speed or direction) across a layer, mountainous influences need to be accounted for.	Shallow layer of high relative humidity (>= 70%) with temps from 0C to -20C.
2	Weak or moderate system exists with strong potential to develop further	Current RADAR reflectivities in the 30-40 DBz range (or lower)	Significant wind shear (speed or direction) across a layer, mountainous influences need to be accounted for.	Moderate layer of high relative humidity (>= 70%) with temps from 0C to -20C. (current or forecast)
3	Moderate to strong system already exists or development of strong system imminent/likely	Current RADAR reflectivities in the 40-50 DBz range (or lower)	Large wind shear (speed or direction) across a layer, mountainous influences need to be accounted for.	Moderate to deep layer of high relative humidity (>= 80%) with temps from 0C to -20C. (current or forecast)
4	Extremely strong system, intensity could remain as is over the forecast period	Current RADAR reflectivities >50 DBz	Very large wind shear (speed or direction) across a layer, mountainous influences need to be accounted for.	Deep layer of high relative humidity (>= 90%) with temps from 0C to -20C. (current or forecast)

D.7 Training: Guide to the Baseline AWIN Tool

In this experiment, you will conduct your tasks in two conditions, *Concept-A* and *Concept-B*. You will use two different versions of the AWIN flight planning software. This section will train you on the features common to both versions of the AWIN software. After being trained in the baseline features, you will receive training on features specific to *Concept-A* and *Concept-B* immediately prior to the experimental trials.

D.7.1 Overview

The tool is structured so that it is very similar to other computer programs that you may already use. When you first open the tool there are two panels and four different windows that are visible, as seen in Figure 24:

1. The Menu Bar
2. The Top Button Pane
3. The World Map
4. Vertical View
5. A Side Control Panel, that is divided into two.

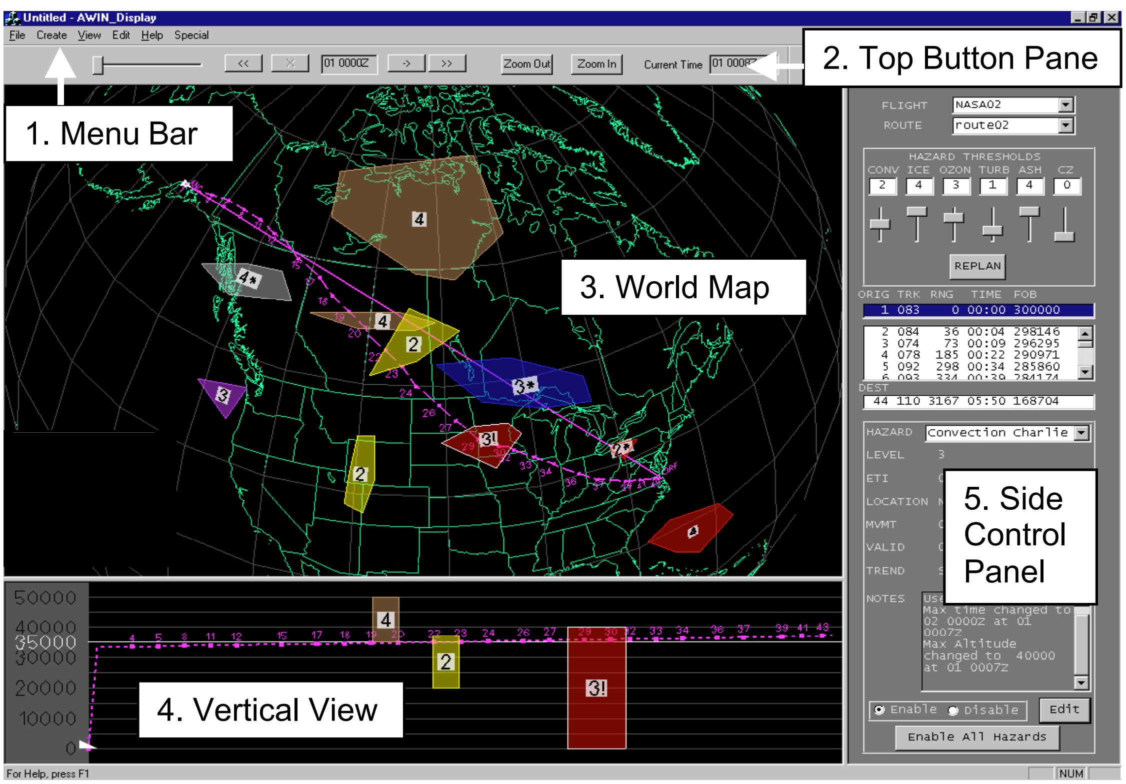


Figure 24. AWIN Interface.

D.7.2 Feature Overview

Our descriptions of the different Windows and Panels in the tool will be clearer if the following concepts are introduced first.

The flight planning tool distinguishes 4 routes.

- The **focus route** is the route of current interest. Additional information is displayed for the focus route. This information is displayed in the lower side panel.
- The **reference route** is the baseline for comparing a route to the list of route alternatives. This can be changed in the Route Comparison Tool.
- The **active route** is the route that has been selected as the candidate route for planning purposes. The dispatcher can request to see his flights and the system responds by showing each flight and its active route. Only one active route is specified per flight. The active route is green in the World View and Vertical View panels.
- The **hover route** is a transient route selection (indicated by placing the mouse over a route path in the world display). A hover route displays additional information about the route in a temporary pop-up window.

Keeping this in mind, we can now look at the different Windows and Panels that are available in the tool.

D.7.3 Menu Bar

The menu bar has 7 different menus. Only two of them will be used.

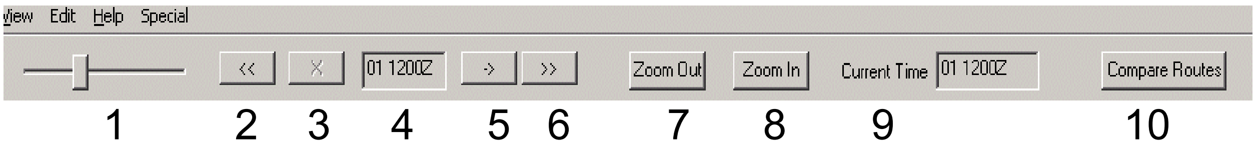
1. **View**: This menu allows you to change what items are displayed on the map, such as airports or winds, and for the tool, such as which windows are visible. There are several View menu options:
    - **Airports.** Display airport labels.
    - **Wind.** Display a wind grid on the Map View *for the selected altitude*. You can select the altitude on the Vertical View by clicking on an altitude on the left vertical axis. You can also change the altitude up or down using the View | Altitude menu option (see below).
    - **Waypoint Labels.** Display waypoint labels on the Map View.
    - **Zoom.** Zoom in or out of the Map View.
    - **Rotate Map.** Rotate the Map View.
    - **Altitude.** Change the altitude setting of the Vertical View (and alos the wind grid on he Map display iif the View | Wind menus option is selected (see above)). You can also change th altitude via the keyboard: up (“a”) or down (“A”).
    - **Comparison Tool:** Brings up separate window that displays all the routes with information for comparison (discussed later). This allows you to compare the different routes that are available for a given flight on factors such as estimated time and fuel use.

2. **Animate**: This menu allows the user to see what will occur during a route. For example, **Start** will begin the animation and show the user the positions of the plane and winds during the course of a given route.

3. The menus **File, Create, Edit, Help** and **Special** will not be used during this experiment.

D.7.4 Top Button Pane

This pane is similar to the Animate menu on the menu bar. It allows you to control the animation of a route.



1. **Time slider**. Valid 12 hours. You can slide the button forward or back to scroll through time.
2. **Rewind fast**.
3. **Stop**.
4. **Movie player time**. The time of picture you are seeing in World View.
5. **Forward time slow**.
6. **Fast Forward time**.
7. **Zoom out**. Decreases scale on the World View.
8. **Zoom in**. Increases scale on the World View.
9. **Current time**. Simulated clock-on-the-wall time.
10. **Compare routes**. This option has the same effect as the Comparison Tool option in the View menu. This button is only enabled if a flight has been selected. Difference calculations are based on the reference route.

D.7.5 The World View

The World View shows the lateral course of flights superimposed on a global world map. The World View has the following features.

1. **Time**: Time of the picture you are seeing. Same as movie player time.
2. **Routes**: All of the different routes for a given flight will be displayed. When you select a route, either by clicking on it or by selecting it using the Side Control Panel (see below), you make it Focused. This Focus route will be highlighted in magenta. Details of the Focus Route will be in the side control panel. To change the focus route, simply click on a different one.
3. **Keyboard Controls:** You can use your keyboard to change the display in the World View. To Navigate around the globe, use the arrow keys. To zoom in use "z" and to zoom out use "Z".

D.7.6 Side Control Panel

The side control panel is divided into two sections. The top section allows you to:

1. **Choose flights**: By clicking on the arrow next to the word FLIGHT you can open a drop down menu that displays a list of flights.

2. **Choose routes for a given flight**: By clicking on the arrow next to the word ROUTE you can open a drop down menu that displays a list of routes that are available for the selected flight.
3. **See the waypoint details for the focus route**: This lists the details of the route, by waypoint, time, and destination.

The bottom section, as mentioned above, will display the details for the focus route.

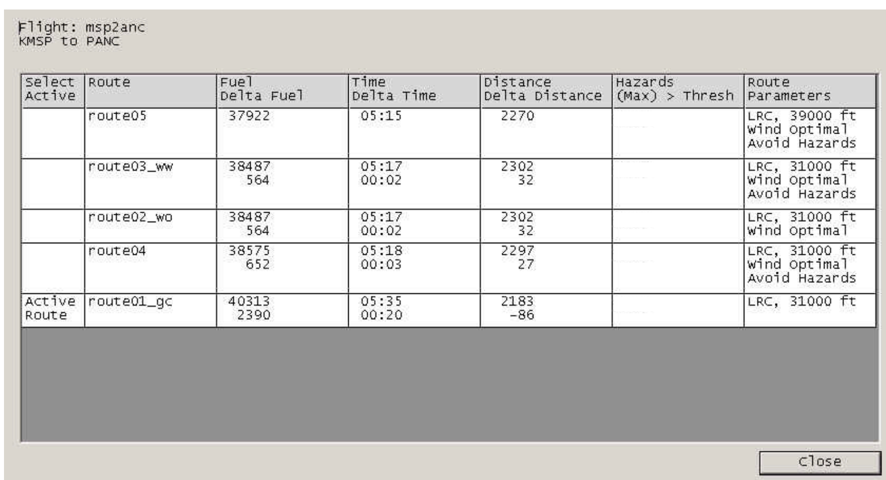
D.7.7 Vertical View

This view is located underneath the World Map and shows the vertical track of the selected flight plan. It shows the selected route (the focus route).

Selected Altitude. The left vertical axis displays the selected altitude. You can display a wind grid on the Map View *for this selected altitude*. You can change the altitude on the Vertical View by clicking on an altitude on the left vertical axis. You can also change the altitude up or down using the View | Altitude menu option (see below). Finally, you can also change the altitude via the keyboard: up ("a") or down ("A").

D.7.8 Route Comparison

One of the key features of the tool is its ability to compare the relative merits of different routes for any given flight. This is done using the **Route Comparison Tool**, as seen in Figure 25.



Flight: msp2anc
KMSP to PANC

Select Active	Route	Fuel Delta Fuel	Time Delta Time	Distance Delta Distance	Hazards (Max) > Thresh	Route Parameters
	route05	37922	05:15	2270		LRC, 39000 ft Wind Optimal Avoid Hazards
	route03_ww	38487 564	05:17 00:02	2302 32		LRC, 31000 ft Wind Optimal Avoid Hazards
	route02_wo	38487 564	05:17 00:02	2302 32		LRC, 31000 ft Wind Optimal
	route04	38575 652	05:18 00:03	2297 27		LRC, 31000 ft Wind Optimal Avoid Hazards
Active Route	route01_gc	40313 2390	05:35 00:20	2183 -86		LRC, 31000 ft

Close

Figure 25. Route Comparison Tool.

There are seven columns:

1. **Select Active**. Clicking on a particular row in this column will designate that route as the active route (as defined on page 94).
2. **Route**. The name of the route.
3. **Fuel / Delta Fuel**. The first number is the total fuel use (in pounds) for the route. The second number (directly underneath) is the difference (in pounds) between the fuel use for the route and the fuel use in the reference route.

4. **Time / Delta Time**. The first number is the total flying time (hours:minutes) of the route. The second number (directly underneath) is the difference (in hours:minutes) between the flying time of the route and the flying time of the reference route.
5. **Distance / Delta Distance.** The first number is the total distance (miles) of the route. The second number (directly underneath) is the difference (in miles) between the distance of the route and the distance of the reference route.
6. **Hazards (Max) > Thresh.** Not applicable in the baseline.
7. **Route Parameters.** Attributes of the route.

D.7.9 Tasks

The tool allows you to view flight routes, compare flight parameters (fuel burn, flying time, etc).

During the next six trials, you will be asked to follow the general instructions on page 90, using the tool to make routing decisions, and MOCK to assess the impact weather may have on that routing decision.

The following sub-sections describe how to accomplish tasks that support your instructions.

Loading weather and route data

The flight planning system will be set up with a map view of North America. For each trial, you will need to load the correct data in the flight planning tool and the MOCK weather display.

Tool	Action
Flight Planning Tool	Select the flight that corresponds to this scenario's city-pair.
Flight Planning Tool	The screen should now display a set of company routes and the auto-generated route between a city pair in the USA
MOCK	Select the date on the left side frame that corresponds to the date of the scenario.
MOCK	The MOCK display will allow you to page through the weathers displayed to assess the impact of that weather on the company flights

Select a flight

In order to see more information about a route, you can select a flight.

Tool	Action
Flight Planning Tool	Using the top portion of the **Side Control Panel.**

View routes

The Flight Planning Tool displays all routes.

Tool	**Action**
Flight Planning Tool	Routes are displayed on the **World View** and in the **Vertical View**
Flight Planning Tool	To see the details for any routes, click on it, and the details will appear in the bottom portion of the **Side Control Panel.**
MOCK	Weather information for the trial can be found on the MOCK page with the corresponding date. All the weather information found in MOCK is valid for the routes found in the tool.

Compare routes

Using the Flight Planning Tool you can view the relative strengths and weaknesses for the different routes.

Tool	**Action**
Flight Planning Tool	Use the **Compare Routes** button on the far right side of the **Top Button Pane,** or
Flight Planning Tool	Select **Comparison Tool** from the **View Menu**

Select an active route

There are two ways of selecting the active route.

Tool	**Action**
Flight Planning Tool	Click on the route in the **World View** to make it focused, and then click the Make Active button visible in the lower portion of the **Side Control Panel,** or
Flight Planning Tool	Click the **Make Active** button in the **Comparison Tool**

Final route designation

In order to end the trial, you will need to make a final selection of the route you wish to dispatch.

Tool	**Action**
Flight Planning Tool	Designate the route you wish to select as the “Active Route”

D.8 Training: Concept A

In the proposed *Concept-A* system, hazard information in the form of polygons will be integrated with route information on the same display. Dispatchers will be able to view both vertical and lateral route information. All company routes will be available in the tool. In addition, an auto-generated route will be available that optimizes fuel while avoiding hazardous weather. Fuel information will be provided with the routes. Dispatchers can use the fuel information for route comparisons.

A company meteorologist was responsible for generating the hazard polygons, and their severity levels, using the images and weather information found in MOCK. Company policy (encapsulated in the weather training of page) was followed when the entering weather hazards and defining severities.

Under Concept A, there are some additional features available and some changes over the baseline description found in Section (page 94). ***The tool in the Concept-A condition is the sum of the baseline training and this concept-A training.***

D.8.1 The Menu Bar

In *Concept-A*, the menu Bar has the following additional features:

1. **View**. This menu allows you to change which items are displayed on the map, such as hazards.
2. **Animate**. Hazard motion is included in the animation.

D.8.2 World View

In *Concept-A*, hazard information is integrated with the World View, which has the following additional features:

1. **Hazards**: All of the hazards at a given altitude will be viewable initially. You can choose to "hide" different types of hazards (See 2). Hazards can also be Focused. To make Focus a hazard, click on it. The Focus Hazard will have a white outline around it. Details of this hazard will be available in the side control panel. You may select between hazards by clicking on them.
2. **Hazard palette.** The Hazard palette is initially located at the bottom left corner of the World View (see Figure 1). You may control which groups of hazards are displayed by checking them on (displayed) or off (not displayed). All hazards start off checked. This palette can be repositioned.

D.8.3 Menu Bar

The menu bar has 7 different menus. Inclusion of hazard information results in the following additions to the baseline:

1. **Animate**: This menu allows the user to see what will occur during a route. For example, **Start** will begin the animation and show the user the positions of the plane, hazards, and winds during the course of a given route.

D.8.4 Side Control Panel

Concept-A includes hazards, and as such the side control panel has additional controls to:

1. **Set the thresholds for the hazard**: This can be done by changing the threshold number below the type of hazard or by using the slider. The airline meteorologist has set these thresholds, thus dispatchers will **not** be allowed to change them.

The bottom section, as mentioned above, will display the details for the focus route or for the focus hazard.

D.8.5 Vertical View

In *Concept-A*, the Vertical View has the following additional features:

1. **Route**: The selected route (the focus route).
2. **Hazards**: Any hazards that the route will encounter are shown as well. Unlike the World view, only the hazards that the route intersects will be shown.

D.8.6 Route Comparison Tool

In *Concept-A* version of the flight planning tool, the Route Comparison Tool has an additional information item:

1. **Hazards (Max) > Thresh.** Lists all the hazards the route penetrates where the hazard severity is above threshold.

D.8.7 Tasks

The tool allows you to view flight routes, compare flight parameters (fuel burn, flying time, etc), and view weather hazard information.

During the next six trials, you will be asked to follow the general instructions on page 90, using the tool to make routing decisions, and MOCK to assess the impact weather may have on that routing decision.

The following sub-sections describe any additional tasks in *Concept-A* that support your instructions, and should be considered *in addition* to the tasks described in the Baseline Training of Page 98.

Loading Weather and Route Data

The flight planning system will be set up with a map view of North America. For each trial, you will need to load the correct data in the flight planning tool and the MOCK weather display.

Tool	Action
Flight Planning Tool	Loading the flights will automatically load the routes and hazard information as well.

View Hazards and Routes

The Flight Planning Tool has hazard information encapsulated as polygons. All routes are displayed as well.

Tool	Action
Flight Planning Tool	Weather is displayed as shaded polygons on the **World View** and in the **Vertical View**
Flight Planning Tool	To see the details for any hazards or routes, click on it, and the details will appear in the bottom portion of the **Side Control Panel.**
MOCK	Weather information for the trial can be found on the MOCK page with the corresponding date. All the weather information found in MOCK is valid for the routes found in the tool.

D.9 Training: Concept B

In the proposed *Concept-B* system, dispatchers will be able to view both vertical and lateral route information. All company routes will be available in the tool. In addition, the user can modify an existing route manually to create a route that optimizes fuel while avoiding hazardous weather. Fuel information will be provided with the routes. Dispatchers can use the fuel information for route comparisons.

Current weather information can be found in the images and weather information found in MOCK. Company policy (encapsulated in the weather training of page) should be followed when the deciding which weather is hazardous and should be avoided.

Under Concept B, there are some additional features available and some changes over the baseline description found in Section (page 94). ***The tool in the Concept-B condition is the sum of the baseline training and this concept-B training.***

D.9.1 Modify an Existing Route

In *Concept-B*, the user can modifying an existing route.

1. Decide which route you would like to modify.
2. Select the route on the **Side Control Panel**, as shown in Figure 26.

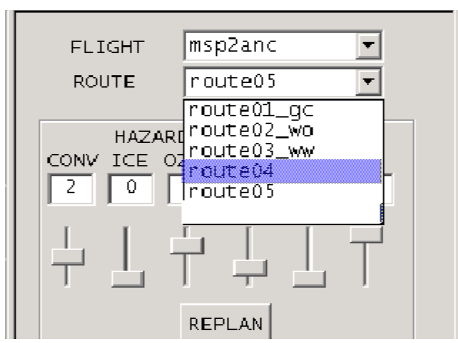


Figure 26. Selecting a route on the Side Control Panel.

3. On the **Edit Menu**, select **Route** or click the **Edit** button on the lower Side Control Panel.
4. A new window will appear. This is the **Flight and Route Planning** window, shown in Figure 5. Much of the pre-existing flight information will already be filled in for you.
5. Please note that the AWIN tool has *automatically* assigned this route a new unique name.

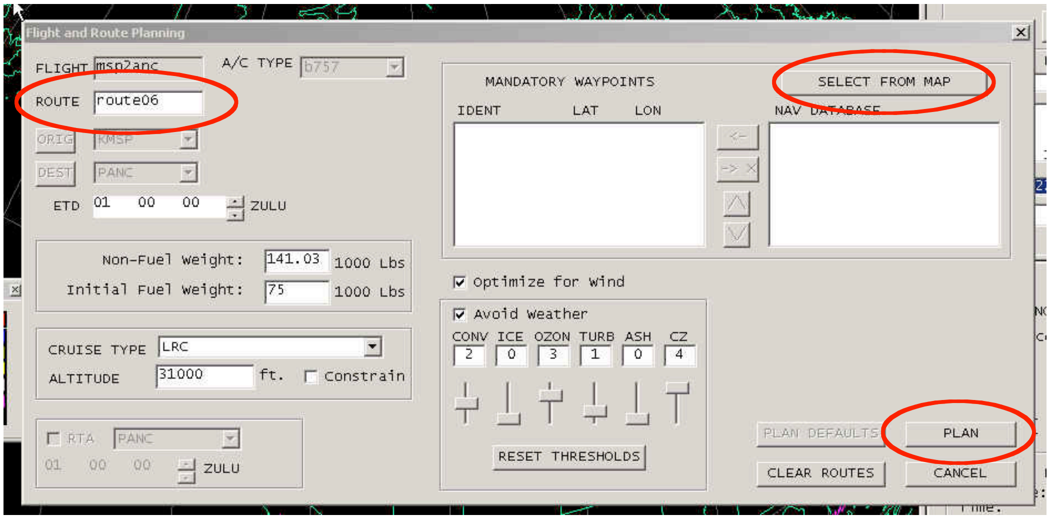


Figure 27. Flight and Route Planning window.

6. Rename this route as "*RouteXX-modified*" where "*RouteXX*" was the original route you chose to modify. If you are modifying an already-modified route, change the name back to the modified route name.
7. You can insert new mandatory waypoints for the route as a way of modifying where the route goes. If you do not want to select mandatory waypoints for the route, skip to Step 8.
   a. Click on the **Select from Map** button in the upper right corner of the window.
   b. This will take you back to the World Map. Click on the map where you would like a waypoint to be.
   c. This will take you back to the **Flight and Route planner**. Your selected waypoint should now appear in the list of mandatory waypoints. To select more waypoints, click on the **Select from Map** button again.
   d. You can remove a waypoint from the list. Highlight the waypoint (see above) and then click on the button "-> x". To move a waypoint up or down in the list, click on the up or down buttons.
8. You can also select a different cruise altitude for the flight by typing in the desired cruise altitude in the **ALTITUDE** box.
9. When you are finished selecting waypoints for your route, click on the **Plan** button in the lower right hand corner of the **Flight and Route Planner**.
10. This will bring up a **Save As** Window. Click **Save** if the file has been renamed.

D.10 weather familiarization session

There were four weather familiarization sessions, one for each weather case (A, B, C, D). Below are the directions for each case, immediately followed by three experimental trials.

D.10.1 Briefing Guide Directions

In the weather familiarization session you will be given dedicated amount of time (5 minutes) to become familiar with the weather via the graphics on the MOCK display. You are **not** required to look at every link. The goal of the familiarization period is to allow you a dedicated amount of time to browse the data (weather graphics) and gain a general appreciation of the weather situation. Once the dedicated weather familiarization session is over, you will begin three trials where that weather data is valid. You will continue to have access to the weather data during the trials themselves.

Please start now by clicking on the MOCK leftmost menu item for **Case X.**

D.10.2 Situation Awareness Question

Immediately after the Weather Familiarization session was complete, subjects were asked to answer a question without referring back to MOCK. This was to test the situation awareness they had built up during the weather familiarization session. Each weather case had a specific question, as follows in Table 59.

Table 59. Situation Awareness probes.

Weather Case	Question
A	Is it raining in the Bismarck, North Dakota area?
B	Assuming a cruise altitude of 28,000 feet, would a route passing through the Minneapolis-St. Paul area in 3 hours encounter significant wind shear (speed or direction)?
C	Assuming a cruise altitude of 23,000 feet, would a route passing over the San Francisco area in 4 hours encounter a moderate layer of high relative humidity (>= 70%) with temperatures from 0 to –20C?
D	Are Dallas/Fort Worth area RADAR reflectivities currently in the 40-50 DBz range (or lower)?

D.11 Sample Trial Sheet

6.4.1 Scenario Instructions

It is your job to choose the most fuel-efficient route that avoids any hazardous weather. The assessment of weather as hazardous is governed by the rules defined by the meteorologist on page 11.

Weather information for this scenario can be found in the MOCK Dispatcher briefing page by selecting the correct date (see below) from the menu.

6.4.2 Scenario Information

Data Type	**Value**
Case	A
Date	June 06, 2001
City-Pair:	PHX – ORD
	Phoenix Sky Harbor International Airport – Chicago O'Hare International Airport
Scheduled Departure Time	0215 Z
Current Time:	0015 Z

6.4.3 Aircraft Parameters:

Parameter	**Value**
Aircraft Type	737-400
Cruise Type	M80 (Fixed Mach = .80)
Initial Fuel Weight	36,000 lbs.
Non-Fuel Weight	78,000 lbs.
Recommended Cruise Altitude	29,000 ft.

6.4.4 Selected Route

Please fill in the following information about the route you selected for dispatch.

Parameter	**Value**
Route Name	
Fuel Consumption	lbs.

D.12 Post-Block Questionnaire, Concept-A

1. I feel that I had a good awareness of the weather situations.

Strongly Disagree		Disagree		Agree		Strongly agree	
1	2	3	4	5	6	7	N/A

Please Explain:

2. I had all the information I needed to make a good routing decision.

Strongly Disagree		Disagree		Agree		Strongly agree	
1	2	3	4	5	6	7	N/A

Please Explain:

If not, what information was missing?

3. I feel confident that the hazards accurately captured the boundary and severity of the weather hazards.

Strongly Disagree		Disagree		Agree		Strongly agree	
1	2	3	4	5	6	7	N/A

Please Explain:

4. I am confident that I selected fuel-efficient routes that avoided hazardous weather.

Strongly Disagree		Disagree		Agree		Strongly agree	
1	2	3	4	5	6	7	N/A

Please Explain:

5. I feel that the routes I selected were safe.

Strongly Disagree		Disagree		Agree		Strongly agree	
1	2	3	4	5	6	7	N/A

Please Explain:

D.13 Post-Block Questionnaire, Concept-B

1. I feel that I had a good awareness of the weather situations.

Strongly Disagree		Disagree		Agree		Strongly agree	
1	2	3	4	5	6	7	N/A

Please Explain:

2. I had all the information I needed to make a good routing decision.

Strongly Disagree		Disagree		Agree		Strongly agree	
1	2	3	4	5	6	7	N/A

Please Explain:

If not, what information was missing?

3. I am confident that I selected/constructed fuel-efficient routes that avoid hazardous weather?

Strongly Disagree		Disagree		Agree		Strongly agree	
1	2	3	4	5	6	7	N/A

Please Explain:

4. I feel that the routes I selected were safe.

Strongly Disagree		Disagree		Agree		Strongly agree	
1	2	3	4	5	6	7	N/A

Please Explain:

D.14 Post Experiment Questionnaire

D.14.1 Part One. Comparison of Flight Planning Systems

Below you will find a list of statements regarding the two experimental conditions (*Concept-A* and *Concept-B*) of the evaluation. For each statement, use the scale provided to place an **A** marker for the *Concept-A* condition and an **B** marker for the *Concept-B* condition on the line near the number that most closely represents your opinion of the flight planning system used in conjunction with the weather information sources. Please write comments to explain your ratings, in the space below the question.

A – Concept-A **B – Concept-B**

***Example:* I could effectively complete my tasks.**	Strongly Disagree	◄—A—B—► 1 2 3 4 5 6 7	Strongly Agree
1. How well does the tool enhance safety?	Extremely Poorly	◄————► 1 2 3 4 5 6 7	Extremely Well

Please Explain:

2. How well does the tool assist you in comparing flight plans to assess fuel efficiency.	Extremely Poorly	◄————► 1 2 3 4 5 6 7	Extremely Well

Please Explain:

3. How well does the tool assist you in detecting when a flight plan intersects hazardous weather?	Extremely Poorly	◄————► 1 2 3 4 5 6 7	Extremely Well

Please Explain:

4. How difficult was it to read and understand the information on the tool's display?

Extremely Difficult	1	2	3	4	5	6	7	Extremely Easy

Please Explain:

5. How well where you able to assess the severity of weather.

Extremely Poorly	1	2	3	4	5	6	7	Extremely Well

Please Explain:

6. How comfortable are you in ability to use this tool for flight routing.

Extremely Uncomfortable	1	2	3	4	5	6	7	Extremely Comfortable

Please Explain:

7. How valid were the scenarios?

Incomplete, inconsistent with real operations	1	2	3	4	5	6	7	Complete, consistent with real operations

Please Explain:

D.14.2 Part Two. Trust Issues

We are interested in your judgement of how reliable and trustworthy you believe the flight planning tools to be, so let's discuss the notion of trust.

If you think about your trust in people, you probably trust some people more than other, with some people you trust very much, and some you trust very little. We do not trust people equally, and we can express how much we trust a particular person.

Likewise, we can think about trusting things, such as products. For example, I trust my Ford to start in the morning because it has never failed to do so, but I trust my Chevrolet much less since it has a history of trouble. So we can assess our level of trust on a scale from 1 to 10, where:

- 10-I trust this tool as much as the most reliable tool(s) I have used on the job and would definitely rely on it.
- (9,8,7,6) 5-I do not especially mistrust or trust this tool; I believe it may be reliable or unreliable
- (4,3,2,1) 0-I do not trust this tool at all and would never want to rely on it.

For example, you can rate your trust, your judgement of predictability, your judgement of dependability and your faith in questions like the following:

1. The local bus service to be on time.
2. Your calculator to produce the right answer.
3. Your heating/cooling system to keep you comfortable.
4. Your watch to tell the correct time.

Given this type of rating, please answer the following questions. There are no "right" answers – we are interested in your assessment.

8. Overall, how do you rate your trust that the CONCEPT A flight planner will produce the most fuel-efficient route?

Not At All ◄——► Completely
1 2 3 4 5 6 7 8 9 10

9. How do you rate your trust that the hazard polygons accurately represent the presence of weather hazards (*i.e.,* any significant weather that exists will be represented by hazard polygons within CONCEPT A)?

Not At All ◄——► Completely
1 2 3 4 5 6 7 8 9 10

10. How do you rate your trust that the hazard polygons accurately represent the spatial extent of the weather (*i.e.,* 3D polygon boundaries encompass significant weather)?

Not At All ◄——► Completely
1 2 3 4 5 6 7 8 9 10

11. How do you rate your trust that the hazard polygon severity accurately represents the severity of the weather?

Not At All ← → Completely

1 2 3 4 5 6 7 8 9 10

12. Using the CONCEPT A Tool, how high is your self-confidence in selecting a fuel-efficient route that avoids hazardous weather?

Not At All ← → Completely

1 2 3 4 5 6 7 8 9 10

13. Using the Concept-B Tool, how high is your self-confidence in constructing (via the manual manipulation of waypoints on a company route) a fuel-efficient route that avoids hazardous weather?

Not At All ← → Completely

1 2 3 4 5 6 7 8 9 10

14. Rate the trust you had in the weather information used in the trials?

Not At All ← → Completely

1 2 3 4 5 6 7 8 9 10

15. Please take a few moments to elaborate on any other trust issues you may have?

D.14.3 Part Three. Open-Ended Questions

16. Please list best and worst features of any current flight planning systems with which you are familiar.

Best:

Worst:

17. Please comment on the compatibility of the Honeywell prototype system with current systems. Is there anything you would need to "unlearn?"

18. How does the reliability of weather information affect your willingness to plan a route through a hazard?

19. Are there rules of thumb you use in terms of when to consider a weather prediction certain and plan accordingly, and when you consider it uncertain and essentially ignore it?

20. Please list best and worst features for flight planning with the Honeywell prototype flight planning system.

Best:

Worst:

21. Do you have any suggestions to make flight planning with the Honeywell prototype easier? (*e.g.,* additional functionality/ information to incorporate, better display formats, tools to help you make decisions with this data, etc.)

22. What is your perception of the validity of the experimental equipment, information sources, and scenarios, and efficacy of training.

23. Would you want to use this tool? Why/Why not”

24. The TLX Workload Scale that you used has six dimensions:

- **Mental Demand** – how much mental effort is required to perform the task (*e.g.,* thinking, deciding, remembering)
- **Physical Demand** – how much physical effort is required to perform the task (*e.g.,* pushing, pulling, reaching, stretching)
- **Temporal Demand** – how much time pressure you feel to complete the task (*e.g.,* relaxed pace or fast and furious?)
- **Performance –** how successful you feel you are in completing the task
- **Effort** – how hard you work to complete the task
- **Frustration** – how aggravated or annoyed versus secure or content you feel about accomplishing the task.

Do you feel you were overly taxed on any of these dimensions? If so, please explain.

D.15 Experiment Purpose Explanation Letter

Dear Participant:

The experiment in which you have just participated, supports a research effort for improved aviation weather information that is jointly sponsored by the Honeywell Technology Center and the NASA Safety Program's Aviation Weather Information (AWIN) project through a cooperative agreement.

The NASA Safety Program was initiated as a response to the 1997 presidential goal of reducing fatal accident rate for aviation by 80% within 10 years. Recognizing that weather continues to be identified as a causal factor in about 30% of all aviation accidents, the NASA AWIN project was established to help meet this safety goal.

NASA AWIN aims to provide with the aid of industry, academia, and aviation user communities, more accurate, timely, and usable information to pilots, dispatchers, and air traffic controllers. With this improved aviation weather information, we hope to facilitate detection and avoidance of weather hazards, to provide these users with better strategies when encountering weather hazards, and ultimately make aviation safer.

NASA AWIN is lead by Mr. Paul Stough at the NASA Langley Research Center in Hampton, Virginia. More information about the NASA AWIN project is available on the project website: http://awin.larc.nasa.gov.

We greatly appreciate your participation in this experiment, and thereby, your assistance in making aviation safer through improved weather information. If you are interested in further assisting NASA AWIN efforts, please contact us via the website above or, if you do not have internet access, please contact me directly: by phone at 757-864-2030, by email at k.a.latorella@larc.nasa.gov, or by postal mail at Crew/Vehicle Integration Branch, MS 152, NASA Langley Research Center, Hampton VA 23681-0001.

Sincerely,
Kara Latorella. Ph.D.
NASA Langley Research Center

NOTE: Essential to the validity of the results is that each subject does not have full foreknowledge of the manipulations contained within this experiment. Therefore, we ask you not to relay any of the details of the experiment to those who might later participate in this study.

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.			
1. AGENCY USE ONLY (*Leave blank*)	**2. REPORT DATE** April 2003	**3. REPORT TYPE AND DATES COVERED** Technical Memorandum	
4. TITLE AND SUBTITLE Evaluation of a Dispatcher's Route Optimization Decision Aid to Avoid Aviation Weather Hazards			**5. FUNDING NUMBERS** 728-40-10-03
6. AUTHOR(S) Michael C. Dorneich, Olu Olofinboba, Steve Pratt, Dannielle Osborne, Thea Feyereisen, and Kara Latorella			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) NASA Langley Research Center Hampton, VA 23681-2199			**8. PERFORMING ORGANIZATION REPORT NUMBER** L-18271
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Langley Research Center Hampton, VA 23681-2199			**10. SPONSORING/MONITORING AGENCY REPORT NUMBER** NASA/TM-2003-212172
11. SUPPLEMENTARY NOTES			
12a. DISTRIBUTION/AVAILABILITY STATEMENT Unclassified-Unlimited Subject Category 03 Distribution: Nonstandard Availability: NASA CASI (301) 621-0390			**12b. DISTRIBUTION CODE**
13. ABSTRACT *(Maximum 200 words)* This document describes the results and analysis of the formal evaluation plan for the Honeywell software tool developed under the NASA AWIN "Weather Avoidance using Route Optimization as a Decision Aid" project. The software tool aims to provide airline dispatchers with a decision aid for selecting optimal routes that avoid weather and other hazards. The evaluation assessed gains in safety, in fuel efficiency of planned routes, and in time efficiency in the pre-flight dispatch process through the use of the AWIN decision aid.			
14. SUBJECT TERMS Weather avoidance, route optimization, decision aid, aviation, dispatch, AOC			**15. NUMBER OF PAGES** 121
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UL